(4)

AD-A215 168

TECHNICAL REPORT BRL-TR-3061

BRL

A SMALL SAMPLE POWER STUDY OF THE ANDERSON-DARLING STATISTIC AND A COMPARISON WITH THE KOLMOGOROV AND THE CRAMÉR-VON MISES STATISTICS

LINDA L. CRAWFORD MOSS
MALCOLM S. TAYLOR
HENRY B. TINGEY

DECEMBER 1989

U.S. ARMY LABORATORY COMMAND

BALLISTIC RESEARCH LABORATORY
ABERDEEN PROVING GROUND, MARYLAND

89 11 22 067

DESTRUCTION NOTICE

Destroy this report when it is no longer needed. DO NOT return it to the originator.

Additional copies of this report may be obtained from the National Technical Information Service, U.S. Department of Commerce, Springfield, VA 22161.

The findings of this report are not to be construed as an official Department of the Army position, unless so designated by other authorized documents.

The use of trade names or manufacturers' names in this report does not constitute indorsement of any commerical product.

UNCLASSIFIED

SECURITY CLASSIFICATION OF THIS PAGE

REPORT DOCUMENTATION PAGE

Form Approved
OMB No. 0704-0188

1a. REPORT SECURITY CLASSIFICATION UNCLASSIFIED	1b. RESTRICTIVE MARKINGS
2a. SECURITY CLASSIFICATION AUTHORITY	3. DISTRIBUTION/AVAILABILITY OF REPORT Approved for public release; distribution unlimited.
2b. DECLASSIFICATION/DOWNGRADING SCHEDULE	
4. PERFORMING ORGANIZATION REPORT NUMBER(S) **BRL-TR-3061**	5. MONITORING ORGANIZATION REPORT NUMBER(S)
6a. NAME OF PERFORMING ORGANIZATION USA Ballistic Research Lab 6b. OFFICE SYMBOL *(If applicable)* SLCBR-SE-P	7a. NAME OF MONITORING ORGANIZATION US Army Ballistic Research Laboratory
6c. ADDRESS *(City, State, and ZIP Code)* Aberdeen Proving Ground, MD 21005-5066	7b. ADDRESS *(City, State, and ZIP Code)* ATTN: SLCBR-DD-S Aberdeen Proving Ground, MD 21005-5066
8a. NAME OF FUNDING/SPONSORING ORGANIZATION 8b. OFFICE SYMBOL *(If applicable)*	9. PROCUREMENT INSTRUMENT IDENTIFICATION NUMBER
8c. ADDRESS *(City, State, and ZIP Code)*	10. SOURCE OF FUNDING NUMBERS PROGRAM ELEMENT NO. / PROJECT NO. / TASK NO. / WORK UNIT ACCESSION NO.

11. TITLE *(Include Security Classification)*
A SMALL SAMPLE POWER STUDY OF THE ANDERSON-DARLING STATISTIC AND A COMPARISON WITH THE KOLMOGOROV AND THE CRAMÉR-VON MISES STATISTICS

12. PERSONAL AUTHOR(S)
Linda L. Crawford Moss, Malcolm S. Taylor, Henry B. Tingey

13a. TYPE OF REPORT	13b. TIME COVERED	14. DATE OF REPORT *(Year, Month, Day)*	15. PAGE COUNT
TR	FROM ______ TO ______		

16. SUPPLEMENTARY NOTATION

17. COSATI CODES			18. SUBJECT TERMS *(Continue on reverse if necessary and identify by block number)*
FIELD	GROUP	SUB-GROUP	Anderson-Darling, Goodness-of-Fit Tests, Power, Kolmogorov, Cramér-von Mises

19. ABSTRACT *(Continue on reverse if necessary and identify by block number)*

The Anderson-Darling goodness-of-fit procedure emphasizes agreement between the data and the hypothesized distribution in the extremes or tails. An improved table of the quantiles of the Anderson-Darling statistic, useful for small sample sizes, was constructed using the Cray-2 supercomputer. The power of the Anderson-Darling test is compared to the Kolmogorov and the Cramér-von Mises tests when the null hypothesis is the normal distribution and the alternative distributions are the Cauchy, the double exponential, and the extreme value distributions.

20. DISTRIBUTION/AVAILABILITY OF ABSTRACT ☐ UNCLASSIFIED/UNLIMITED ☒ SAME AS RPT ☐ DTIC USERS	21. ABSTRACT SECURITY CLASSIFICATION UNCLASSIFIED
22a. NAME OF RESPONSIBLE INDIVIDUAL Linda L. Crawford Moss	22b. TELEPHONE *(Include Area Code)* 278-6658 22c. OFFICE SYMBOL SLCBR-SE-P

DD Form 1473, JUN 86 *Previous editions are obsolete.* SECURITY CLASSIFICATION OF THIS PAGE

UNCLASSIFIED

TABLE OF CONTENTS

Page

LIST OF ILLUSTRATIONS

LIST OF ILLUSTRATIONS (contd)

ACKNOWLEDGEMENTS

The authors thank Nancy H. Dunn, TECOM and David W. Webb, BRL, for their helpful comments after reviewing the report. We also express our appreciation to Annette L. Wiseman for formatting this report.

INTENTIONALLY LEFT BLANK.

I. INTRODUCTION

Consider a random sample $X_1, X_2, \ldots, X_n$ from a population with a continuous distribution function. One method of testing the hypothesis that the n observations come from a population with a specified distribution function F(x) is by a chi-square test. This test requires a subjective partitioning of the real line **R** and a comparision of the empirical histogram with the hypothetical histogram. A more objective method, is to compare the empirical distribution function $F_n(x)$ with the hypothetical distribution function F(x). The empirical distribution function based on n observations is defined as $F_n(x) = \frac{k}{n}$ if exactly k observations are less than or equal to x for $k = 0, 1, \ldots, n$.

To compare the empirical and hypothetical distribution functions a measure of their difference is required. Addressing this, Anderson and Darling [1952] considered the following metrics in function space:

$$W_n^2 = n \int_{-\infty}^{\infty} [F_n(x) - F(x)]^2 \, \psi[F(x)] \, dF(x) \tag{1}$$

and

$$K_n = \sup_{-\infty < x < \infty} \sqrt{n} \, |F_n(x) - F(x)| \sqrt{\psi[F(x)]} \, . \tag{2}$$

Samples producing large values of W_n^2 (or K_n) lead to rejection of the null hypothesis that the population distribution function is F(x). One of the contributions of Anderson and Darling was the incorporation of a non-negative weight function ψ in (1) and (2). By a suitable choice for ψ, specific ranges of values of the random variable X, corresponding to different regions of the distribution F(x), may be emphasized. For $\psi[F(x)] \equiv 1$, W_n^2 becomes the Cramér-von Mises statistic [Cramér, 1928 and von Mises, 1931] and K_n becomes the Kolmogorov statistic [Kolmogorov, 1933].

The tails of the distribution function will be accentuated in the investigation detailed in this paper; Anderson and Darling suggest using

$$\psi[F(x)] = \frac{1}{F(x)\,[1 - F(x)]} \, .$$

With this choice for the weighting function, metric (1) becomes the basis for the Anderson-Darling statistic.

In Section II, the Anderson-Darling test statistic is developed; in Section III, the most accurate tabulation to date of the test statistic is provided. In Section IV, the description and the results of a power study are given in which the Anderson-Darling, the Cramér-von Mises, and the Kolmogorov statistics are compared.

II. THE ANDERSON-DARLING STATISTIC

For a fixed value of the random variable X, say $X = x$, the empirical distribution function $F_n(x)$ is a statistic, since it is a function of the sample values $x_1, x_2, \ldots, x_n$. The distribution of this statistic is established as a lemma.

Lemma (1): If $F_n(x)$ is the empirical distribution function corresponding to a random sample $X_1, X_2, \ldots, X_n$ of size n from a distribution $H(\cdot)$, then for a fixed x, $nF_n(x)$ is distributed binomial$(H(x), n)$.

Proof:
$P(nF_n(x) = k) = P(\text{exactly } k \text{ values } x_i \leq x)$, for $k = 0, 1, \ldots, n$.

Let $Z_i = I_{(-\infty, x]}(X_i)$, where the indicator function I is defined as

$$I_{(-\infty, x]}(X_i) = \begin{cases} 1, & \text{if } -\infty < X_i \leq x \\ 0, & \text{otherwise} . \end{cases}$$

Then ΣZ_i counts the number of sample values $x_i \leq x$.
Here each $Z_i \sim$Bernoulli$(H(x))$, so $\Sigma Z_i \sim$binomial$(H(x), n)$.
Therefore,

$$\begin{aligned} P(nF_n(x) = k) &= P(\text{exactly } k \text{ values } x_i \leq x) \\ &= P(\Sigma Z_i = k) \\ &= \binom{n}{k} H(x)^k \left(1 - H(x)\right)^{n-k} . \blacksquare \end{aligned}$$

From Lemma 1,

$$E[F_n(x)] = \frac{1}{n} E\left[nF_n(x)\right] = H(x)$$

and

$$\mathrm{Var}[F_n(x)] = \frac{1}{n^2} \mathrm{Var}\left[nF_n(x)\right] = \frac{1}{n} H(x)\,[1 - H(x)]\,. \tag{3}$$

To assist in the determination of a suitable weighting function $\psi(\cdot)$, that is, a function that will weight more heavily values in the tails of the distribution $F(x)$ at the expense of values closer to the median, consider the expectation of the squared discrepancy $E\left[F_n(x) - F(x)\right]^2$. It is important to keep in mind that the value x is fixed, so $F(x)$ is a constant, and the expectation is with respect to the random variable $F_n(x)$ whose distribution was established in Lemma 1. Then

$$\begin{aligned} n\,E\left[F_n(x) - F(x)\right]^2 &= n\,E\left[F_n(x) - H(x) + H(x) - F(x)\right]^2 \\ &= n\,E\left[\left\{F_n(x) - H(x)\right\} - \left\{F(x) - H(x)\right\}\right]^2 \end{aligned}$$

which, after algebraic manipulation (Appendix A) yields the variance and bias^2

$$= n\left[\frac{1}{n}\left\{H(x)\{1 - H(x)\}\right\} + \left\{F(x) - H(x)\right\}^2\right]. \tag{4}$$

Under the null hypothesis H_o: $H(x) = F(x) \quad \forall x$, (4) becomes

$$n\,E\left[F_n(x) - F(x)\right]^2 = F(x)[1 - F(x)]\,. \tag{5}$$

Anderson-Darling chose as a weighting function, $\psi[F(x)] = \dfrac{1}{F(x)\,[\,1 - F(x)]}$.

Weighting by the reciprocal of (5) takes into consideration the variance of the statistic $F_n(x)$ and also maintains the objective of accentuating values in the tails of $F(x)$.

With this choice of weighting function and without loss of generality assuming $x_1 \leq x_2 \leq \dots \leq x_n$, let $F(x) = u$, $dF(x) = du$, and $F(x_i) = u_i$. Then the Anderson-Darling test statistic (6) can be rewritten as expression (7) by expansion and integration (Appendix B).

$$W_n^2 = n \int_{-\infty}^{\infty} \frac{[F_n(x) - F(x)]^2}{F(x)\,[1 - F(x)]} dF(x)\,, \tag{6}$$

$$W_n^2 = -n - \frac{1}{n} \sum_{i=1}^{n} \Big[(2i-1)\ln u_i + (2(n-i)+1)\ln(1-u_i)\Big]. \tag{7}$$

III. DISTRIBUTION OF THE ANDERSON-DARLING STATISTIC

The asymptotic distribution of W_n^2 was derived by Anderson and Darling [1952]. Lewis [1961] undertook the tabulation of $F(z; n) = P(W_n^2 \leq z)$ for $n = 1, 2, \dots, 8$ and for incremental values of z over the interval [0.025, 8.000]. Lewis' table entries were computed using a Monte Carlo procedure to generate an empirical approximation $F_m(z;n)$ to the distribution function $F(z;n)$ based on m samples of size n. At that time, computational restrictions essentially limited the accuracy of the table entries to within 0.00326 of the true value.

Following an analogous procedure based on expression (7) and the observation that the U_i are distributed U[0,1] [Feller, 1966], the table appearing in Lewis' paper was recalculated using a Cray-2 supercomputer. Table 1 lists the reconstruction of Lewis' table, now accurate within 0.0005. Again, z ranges from 0.025 to 8.000 and for $n = 1, 2, \dots, 10$. The column labeled "∞" contains the asymptotic values, rounded to four decimal places.

To obtain this increased accuracy, a Kolmogorov-type bound [Conover, 1980] was used to construct a 95% confidence band for the distribution function $F(z;n)$. In general, the width of a $(1-\alpha)100\%$ confidence band is equal to twice the value of the $(1-\alpha)100\%$ quantile of the Kolmogorov statistic $K_m = \sup_{-\infty < x < \infty} \sqrt{m}\,|F_m(x) - F(x)|$, where m is the number of Monte Carlo samples of size n used in the construction of $F_m(x)$. With n fixed, the 95% confidence band can be made arbitrarily small by a suitable choice for

TABLE 1. Quantiles of $F(z;n)$ for $n = 1$ thru 10 and ∞.

z	n = 1	2	3	4	5	6	7	8	9	10	∞
0.0250	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000
0.0500	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000
0.0750	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000
0.1000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000
0.1250	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0001	0.0001	0.0001	0.0001	0.0003
0.1500	0.0000	0.0000	0.0000	0.0000	0.0003	0.0006	0.0008	0.0009	0.0010	0.0010	0.0014
0.1750	0.0000	0.0000	0.0000	0.0014	0.0025	0.0029	0.0033	0.0034	0.0036	0.0037	0.0042
0.2000	0.0000	0.0000	0.0022	0.0065	0.0075	0.0081	0.0085	0.0087	0.0089	0.0090	0.0096
0.2250	0.0000	0.0000	0.0120	0.0152	0.0160	0.0166	0.0170	0.0173	0.0174	0.0176	0.0180
0.2500	0.0000	0.0008	0.0260	0.0274	0.0283	0.0288	0.0291	0.0294	0.0294	0.0296	0.0296
0.2750	0.0000	0.0299	0.0431	0.0431	0.0440	0.0444	0.0445	0.0448	0.0447	0.0449	0.0443
0.3000	0.0000	0.0586	0.0625	0.0621	0.0628	0.0629	0.0629	0.0631	0.0628	0.0629	0.0618
0.3250	0.0000	0.0869	0.0840	0.0838	0.0842	0.0839	0.0838	0.0837	0.0833	0.0834	0.0817
0.3500	0.0000	0.1147	0.1071	0.1078	0.1073	0.1068	0.1065	0.1063	0.1058	0.1056	0.1036
0.3750	0.0000	0.1421	0.1318	0.1330	0.1319	0.1311	0.1305	0.1303	0.1296	0.1295	0.1269
0.4000	0.1166	0.1691	0.1579	0.1591	0.1574	0.1564	0.1555	0.1552	0.1547	0.1543	0.1513
0.4250	0.1949	0.1955	0.1853	0.1856	0.1833	0.1821	0.1812	0.1808	0.1803	0.1797	0.1764
0.4500	0.2484	0.2217	0.2136	0.2120	0.2096	0.2083	0.2072	0.2066	0.2061	0.2056	0.2019
0.4750	0.2913	0.2474	0.2420	0.2384	0.2358	0.2344	0.2331	0.2325	0.2321	0.2315	0.2276
0.5000	0.3278	0.2727	0.2693	0.2646	0.2618	0.2603	0.2589	0.2584	0.2579	0.2573	0.2532
0.5250	0.3598	0.2976	0.2955	0.2902	0.2875	0.2859	0.2844	0.2839	0.2834	0.2828	0.2786
0.5500	0.3885	0.3221	0.3209	0.3154	0.3128	0.3111	0.3095	0.3091	0.3085	0.3079	0.3036
0.5750	0.4145	0.3462	0.3455	0.3398	0.3373	0.3356	0.3341	0.3337	0.3329	0.3325	0.3281
0.6000	0.4385	0.3700	0.3693	0.3637	0.3612	0.3595	0.3580	0.3576	0.3569	0.3564	0.3520
0.6250	0.4607	0.3936	0.3922	0.3870	0.3844	0.3827	0.3812	0.3809	0.3803	0.3798	0.3753
0.6500	0.4814	0.4166	0.4143	0.4096	0.4069	0.4052	0.4038	0.4035	0.4029	0.4024	0.3980
0.6750	0.5007	0.4394	0.4357	0.4314	0.4286	0.4270	0.4256	0.4254	0.4248	0.4244	0.4199
0.7000	0.5188	0.4618	0.4563	0.4524	0.4495	0.4481	0.4467	0.4464	0.4459	0.4456	0.4412

TABLE 1. continued

z	n = 1	2	3	4	5	6	7	8	9	10	∞
0.7500	0.5522	0.5056	0.4953	0.4923	0.4893	0.4879	0.4867	0.4864	0.4859	0.4856	0.4815
0.8000	0.5821	0.5458	0.5317	0.5289	0.5261	0.5249	0.5238	0.5235	0.5231	0.5228	0.5190
0.8500	0.6091	0.5794	0.5652	0.5627	0.5601	0.5590	0.5581	0.5580	0.5574	0.5573	0.5537
0.9000	0.6337	0.6090	0.5962	0.5938	0.5916	0.5905	0.5897	0.5897	0.5891	0.5891	0.5858
0.9500	0.6563	0.6358	0.6247	0.6225	0.6206	0.6196	0.6190	0.6188	0.6184	0.6183	0.6154
1.0000	0.6771	0.6601	0.6511	0.6489	0.6473	0.6464	0.6458	0.6457	0.6453	0.6454	0.6427
1.0500	0.6963	0.6823	0.6754	0.6732	0.6718	0.6712	0.6706	0.6704	0.6701	0.6703	0.6680
1.1000	0.7141	0.7028	0.6977	0.6958	0.6944	0.6939	0.6934	0.6933	0.6931	0.6932	0.6912
1.1500	0.7306	0.7217	0.7182	0.7164	0.7153	0.7149	0.7145	0.7144	0.7142	0.7143	0.7127
1.2000	0.7459	0.7392	0.7367	0.7356	0.7345	0.7342	0.7339	0.7339	0.7337	0.7338	0.7324
1.2500	0.7603	0.7555	0.7538	0.7531	0.7522	0.7520	0.7518	0.7518	0.7517	0.7518	0.7508
1.3000	0.7737	0.7704	0.7696	0.7693	0.7686	0.7684	0.7683	0.7683	0.7684	0.7684	0.7677
1.3500	0.7862	0.7844	0.7842	0.7844	0.7838	0.7837	0.7836	0.7836	0.7837	0.7838	0.7833
1.4000	0.7980	0.7975	0.7977	0.7982	0.7978	0.7977	0.7978	0.7977	0.7979	0.7980	0.7978
1.4500	0.8090	0.8097	0.8103	0.8110	0.8108	0.8109	0.8109	0.8108	0.8110	0.8112	0.8111
1.5000	0.8194	0.8210	0.8221	0.8229	0.8228	0.8230	0.8230	0.8230	0.8232	0.8235	0.8235
1.5500	0.8291	0.8316	0.8330	0.8340	0.8340	0.8342	0.8343	0.8343	0.8345	0.8348	0.8350
1.6000	0.8382	0.8416	0.8431	0.8442	0.8444	0.8447	0.8448	0.8448	0.8450	0.8453	0.8457
1.6500	0.8469	0.8508	0.8526	0.8538	0.8541	0.8544	0.8546	0.8546	0.8549	0.8550	0.8556
1.7000	0.8550	0.8595	0.8615	0.8627	0.8631	0.8634	0.8636	0.8636	0.8640	0.8642	0.8648
1.7500	0.8626	0.8677	0.8698	0.8709	0.8715	0.8718	0.8720	0.8720	0.8725	0.8727	0.8734
1.8000	0.8698	0.8754	0.8775	0.8787	0.8793	0.8797	0.8798	0.8799	0.8803	0.8805	0.8814
1.8500	0.8767	0.8826	0.8847	0.8860	0.8866	0.8870	0.8872	0.8872	0.8877	0.8879	0.8888
1.9000	0.8830	0.8893	0.8915	0.8928	0.8934	0.8938	0.8940	0.8941	0.8945	0.8947	0.8957
1.9500	0.8891	0.8956	0.8979	0.8991	0.8997	0.9001	0.9004	0.9006	0.9008	0.9011	0.9021
2.0000	0.8948	0.9015	0.9038	0.9051	0.9058	0.9061	0.9064	0.9066	0.9068	0.9071	0.9082
2.0500	0.9002	0.9070	0.9093	0.9107	0.9113	0.9117	0.9120	0.9121	0.9124	0.9127	0.9138
2.1000	0.9053	0.9122	0.9146	0.9159	0.9165	0.9169	0.9172	0.9174	0.9177	0.9179	0.9190
2.1500	0.9102	0.9171	0.9195	0.9208	0.9214	0.9218	0.9221	0.9223	0.9226	0.9228	0.9239
2.2000	0.9148	0.9218	0.9241	0.9254	0.9259	0.9264	0.9266	0.9268	0.9272	0.9273	0.9285

TABLE 1. continued

z	n = 1	2	3	4	5	6	7	8	9	10	∞
2.2500	0.9191	0.9262	0.9284	0.9297	0.9302	0.9307	0.9309	0.9311	0.9315	0.9316	0.9328
2.3000	0.9232	0.9303	0.9325	0.9337	0.9343	0.9347	0.9349	0.9351	0.9355	0.9356	0.9368
2.3500	0.9271	0.9342	0.9363	0.9375	0.9381	0.9385	0.9387	0.9389	0.9393	0.9394	0.9405
2.4000	0.9308	0.9378	0.9399	0.9411	0.9416	0.9420	0.9423	0.9425	0.9428	0.9429	0.9441
2.4500	0.9342	0.9412	0.9433	0.9444	0.9449	0.9454	0.9456	0.9458	0.9461	0.9463	0.9474
2.5000	0.9375	0.9445	0.9465	0.9476	0.9480	0.9485	0.9487	0.9489	0.9492	0.9494	0.9504
2.5500	0.9407	0.9475	0.9494	0.9505	0.9510	0.9514	0.9516	0.9519	0.9521	0.9523	0.9534
2.6000	0.9436	0.9504	0.9523	0.9532	0.9537	0.9542	0.9544	0.9546	0.9548	0.9550	0.9561
2.6500	0.9465	0.9531	0.9549	0.9558	0.9563	0.9567	0.9570	0.9572	0.9574	0.9576	0.9586
2.7000	0.9491	0.9556	0.9574	0.9583	0.9588	0.9592	0.9594	0.9596	0.9599	0.9600	0.9610
2.7500	0.9517	0.9580	0.9598	0.9606	0.9611	0.9614	0.9617	0.9619	0.9622	0.9623	0.9633
2.8000	0.9541	0.9602	0.9620	0.9628	0.9633	0.9636	0.9638	0.9640	0.9643	0.9644	0.9654
2.8500	0.9564	0.9624	0.9641	0.9648	0.9653	0.9657	0.9659	0.9661	0.9663	0.9665	0.9674
2.9000	0.9586	0.9644	0.9661	0.9668	0.9672	0.9676	0.9678	0.9680	0.9682	0.9684	0.9692
2.9500	0.9607	0.9663	0.9679	0.9686	0.9690	0.9694	0.9696	0.9698	0.9700	0.9701	0.9710
3.0000	0.9626	0.9681	0.9696	0.9703	0.9708	0.9711	0.9713	0.9714	0.9717	0.9718	0.9726
3.0500	0.9645	0.9697	0.9713	0.9719	0.9724	0.9727	0.9729	0.9730	0.9733	0.9734	0.9742
3.1000	0.9663	0.9713	0.9728	0.9735	0.9739	0.9742	0.9744	0.9745	0.9748	0.9749	0.9756
3.1500	0.9679	0.9729	0.9743	0.9749	0.9753	0.9757	0.9758	0.9760	0.9762	0.9763	0.9770
3.2000	0.9695	0.9743	0.9757	0.9763	0.9767	0.9770	0.9771	0.9773	0.9775	0.9776	0.9783
3.2500	0.9711	0.9756	0.9770	0.9776	0.9780	0.9783	0.9784	0.9786	0.9787	0.9788	0.9795
3.3000	0.9725	0.9769	0.9782	0.9788	0.9792	0.9794	0.9796	0.9797	0.9799	0.9800	0.9807
3.3500	0.9738	0.9781	0.9794	0.9799	0.9803	0.9806	0.9807	0.9809	0.9810	0.9811	0.9818
3.4000	0.9751	0.9792	0.9805	0.9810	0.9814	0.9816	0.9817	0.9819	0.9821	0.9822	0.9828
3.4500	0.9764	0.9803	0.9815	0.9820	0.9824	0.9826	0.9827	0.9829	0.9830	0.9831	0.9837
3.5000	0.9775	0.9814	0.9825	0.9830	0.9833	0.9836	0.9837	0.9839	0.9840	0.9840	0.9846
3.5500	0.9786	0.9823	0.9834	0.9839	0.9842	0.9844	0.9846	0.9847	0.9849	0.9849	0.9855
3.6000	0.9797	0.9833	0.9843	0.9848	0.9851	0.9853	0.9854	0.9856	0.9857	0.9857	0.9863
3.6500	0.9807	0.9841	0.9851	0.9856	0.9859	0.9861	0.9862	0.9864	0.9865	0.9865	0.9870
3.7000	0.9816	0.9849	0.9859	0.9863	0.9867	0.9868	0.9869	0.9871	0.9872	0.9872	0.9878

TABLE 1. continued

z	n: 1	2	3	4	5	6	7	8	9	10	∞
3.7500	0.9825	0.9857	0.9867	0.9871	0.9874	0.9875	0.9877	0.9878	0.9879	0.9879	0.9884
3.8000	0.9834	0.9865	0.9874	0.9878	0.9881	0.9882	0.9883	0.9885	0.9886	0.9886	0.9891
3.8500	0.9842	0.9872	0.9880	0.9884	0.9887	0.9888	0.9889	0.9891	0.9892	0.9892	0.9897
3.9000	0.9850	0.9878	0.9887	0.9890	0.9893	0.9894	0.9895	0.9897	0.9898	0.9898	0.9902
3.9500	0.9858	0.9885	0.9893	0.9896	0.9899	0.9900	0.9901	0.9902	0.9903	0.9903	0.9908
4.0000	0.9864	0.9890	0.9898	0.9902	0.9904	0.9905	0.9906	0.9908	0.9909	0.9908	0.9913
4.0500	0.9871	0.9896	0.9904	0.9907	0.9909	0.9910	0.9911	0.9913	0.9913	0.9913	0.9917
4.1000	0.9877	0.9901	0.9909	0.9912	0.9914	0.9915	0.9916	0.9917	0.9918	0.9918	0.9922
4.1500	0.9883	0.9906	0.9914	0.9917	0.9919	0.9919	0.9920	0.9922	0.9922	0.9922	0.9926
4.2000	0.9889	0.9911	0.9918	0.9921	0.9923	0.9924	0.9925	0.9926	0.9927	0.9927	0.9930
4.2500	0.9895	0.9916	0.9922	0.9925	0.9927	0.9928	0.9929	0.9930	0.9931	0.9931	0.9934
4.3000	0.9900	0.9920	0.9926	0.9929	0.9931	0.9931	0.9932	0.9934	0.9934	0.9934	0.9938
4.3500	0.9905	0.9924	0.9930	0.9933	0.9935	0.9935	0.9936	0.9937	0.9938	0.9938	0.9941
4.4000	0.9910	0.9928	0.9934	0.9936	0.9938	0.9938	0.9939	0.9941	0.9941	0.9941	0.9944
4.5000	0.9918	0.9935	0.9941	0.9943	0.9945	0.9945	0.9946	0.9947	0.9948	0.9947	0.9950
4.6000	0.9926	0.9941	0.9946	0.9949	0.9950	0.9951	0.9951	0.9952	0.9953	0.9953	0.9955
4.7000	0.9933	0.9947	0.9952	0.9954	0.9956	0.9956	0.9956	0.9957	0.9958	0.9958	0.9960
4.8000	0.9939	0.9952	0.9957	0.9959	0.9960	0.9960	0.9961	0.9962	0.9962	0.9962	0.9964
4.9000	0.9945	0.9957	0.9961	0.9963	0.9964	0.9964	0.9965	0.9966	0.9966	0.9966	0.9968
5.0000	0.9950	0.9961	0.9965	0.9967	0.9968	0.9968	0.9968	0.9969	0.9969	0.9969	0.9971
5.5000	0.9970	0.9977	0.9979	0.9980	0.9981	0.9981	0.9982	0.9982	0.9982	0.9982	0.9983
6.0000	0.9982	0.9986	0.9988	0.9988	0.9989	0.9989	0.9989	0.9989	0.9990	0.9990	0.9990
7.0000	0.9993	0.9995	0.9996	0.9996	0.9996	0.9996	0.9996	0.9996	0.9996	0.9996	0.9997
8.0000	0.9997	0.9998	0.9998	0.9999	0.9999	0.9999	0.9999	0.9999	0.9999	0.9999	0.9999

m. The commonly tabled [Miller, 1956] asymptotic approximation for the 95^{th} quantile is $1.358/\sqrt{m}$. However, Harter [1980] suggests using

$$\frac{1.358}{\left[m + \frac{r}{3.5}\right]^{\frac{1}{2}}}, \quad \text{where } r = (m+4)^{\frac{1}{2}} \tag{8}$$

for an improved approximation.

Using approximation (8) to construct a 95% confidence band with the width not exceeding 0.001, the value for m must be at least 7,375,881. In this simulation, m was chosen to be 7.4 million[1].

IV. POWER STUDY

The power of the Anderson-Darling test was compared with two other goodness-of-fit procedures based on the empirical distribution function: the Kolmogorov and the Cramér-von Mises tests. The Kolmogorov statistic introduced in Section 1 as metric (2) with weighting function $\psi\,[F(x)] \equiv 1$ becomes

$$K_n = \sup_{-\infty < x < \infty} \sqrt{n}\,|F_n(x) - F(x)| \,. \tag{9}$$

For an ordered sample $x_1 \leq x_2 \leq \cdots \leq x_n$ and $F(x_i) = u_i$, K_n may be evaluated as $\sqrt{n}D$ where $D = \max(D^+, D^-)$ and

$$D^+ = \max_i \left(\frac{i}{n} - u_i\right),$$

$$D^- = \max_i \left(u_i - \frac{i-1}{n}\right).$$

1 The sample values were obtained via a linear congruential uniform random number generator of the form $X_{k+1} = (aX_k + c)_{\text{mod } q}$, where $a = 273673163155_8$, $c = 13_8$, and $q = 2^{48}$. It has the properties of a 'good' random number generator as suggested by Rubinstein [1981]. In addition, several subsets of the random numbers generated were tested for autocorrelations with lag up to 36. For each subset and each lag, the autocorrelation did not exceed 0.06 in absolute value.

The Cramér-von Mises statistic, defined as

$$W_n^2 = n \int_{-\infty}^{\infty} [F_n(x) - F(x)]^2 \, dF(x)$$

can be reduced to (10) for ease of computation (Appendix C);

$$W_n^2 = \sum_{i=1}^{n} \left[u_i - \frac{2i-1}{2n} \right]^2 + \frac{1}{12n} \quad . \tag{10}$$

In the power study, two cases were considered. Case 1 corresponds to the situation in which the parameters of the hypothesized distribution are completely specified. Case 2 corresponds to the situation in which the parameters are not specified and must be estimated from the sample data.

For both case 1 and 2, the null hypothesis is

H_o: A random sample $X_1, X_2, \ldots, X_n$ comes from a normal population

or

H_o: $F(x) = F_o(x)$, where $F_o(x) \sim N(\mu, \sigma^2)$.

As alternative hypotheses, the Cauchy, double exponential, and extreme value distributions were chosen, each with location parameter the same as the null hypothesis. This provided a heavy-tailed, light-tailed, and skewed distribution, respectively, against which the power of the three goodness-of-fit tests are compared.

The power functions do not exist in closed form; they are approximated empirically via a Monte Carlo simulation. To determine a point on the power curve, a large number of samples of size n was generated from a specific distribution serving as the alternative hypothesis. The number of times that the null hypothesis was rejected at a specific level of significance was recorded. The ratio of the number of rejections, Y, to the total number of samples generated, N, provides an estimate, $\hat{p} = Y/N$, of the probability of rejecting the null hypothesis when it should be rejected (power). The value $\hat{p}$ determines a point on the power curve corresponding to a specific sample size n, a specific significance level α, and a specific alternative hypothesis.

To determine the number of samples of size n required for a sufficiently accurate estimate of $\hat{p}$, a nonparametric technique was employed. Since the counter Y is distributed binomial(p,N) where the parameter p is the true but unknown power, and since an approximate confidence interval for p can be constructed [Conover, 1980] using

$$1-\alpha \simeq P\left\{\frac{Y}{N} - z_{(1-\alpha/2)}\left\{\frac{Y}{N}(1-\frac{Y}{N})/N\right\}^{\frac{1}{2}} < p < \frac{Y}{N} + z_{(1-\alpha/2)}\left\{\frac{Y}{N}(1-\frac{Y}{N})/N\right\}^{\frac{1}{2}}\right\}. \quad (11)$$

samples of size n continued to be generated from the alternative distribution until the confidence interval for p given in (11) was sufficiently small.

The confidence interval coefficient $1-\alpha$ was chosen to be 0.975 and the confidence interval width not to exceed 0.025. From (11), this will occur when the inequality $2 \cdot z_{0.9875}\left\{\frac{Y}{N}(1-\frac{Y}{N})/N\right\}^{\frac{1}{2}} \leq 0.025$ is satisfied. For any value of N, the left-hand side of the inequality is maximum when Y/N is 1/2. Substituting Y/N = 1/2 into the inequality yields $N \leq 8037$ — the largest possible number of samples required. In practice, the value for N will usually be much smaller and will change for each estimate $\hat{p}$. This dynamic scheme was chosen rather than fixing N = 8037 for the entire simulation. A minimum value for N of 100 was imposed to prevent premature termination of the procedure.

1. Case 1: Distribution Parameters Specified.

The power study for case 1 specified the parameters of the hypothesized distribution as N(0,1). The results of the study are summarized in Figures 1 – 12. For each of the three distributions serving as an alternative hypothesis, samples of size n = 5, 10, 15, 20 were chosen for study and, as previously mentioned, the location parameters of both the null and alternative hypotheses coincided. The scale parameter for the alternative hypothesis ranged from 0.025 to 3.000 in increments of 0.025.

The level of significance for the study was 0.05. The critical value for each test was determined from tables in Conover [1980] for the Kolmogorov test, Stephens and Maag [1968] for the Cramér-von Mises test, and Table 1 in Section III of this paper for the Anderson-Darling test.

The Anderson-Darling test demonstrated overall superiority for the sample sizes and hypotheses chosen for this study. This is perhaps to be anticipated in view of the emphasis on agreement in the tails by the Anderson-Darling procedure, but the magnitude of difference over the Kolmogorov and Cramér-von Mises tests is impressive.

The power curves corresponding to n = 10, 15, 20 are distinguished by their characteristic of decreasing to a global minimum before becoming monotonically increasing. An explanation of this feature is suggested by consideration of Figures 13 – 15 in which the distribution functions of the N(0,1) and Cauchy $(0,\varsigma)$ are compared. There it is seen (Figure 14) that corresponding to $\varsigma = 0.50$ the two distribution functions are similar; an increase (decrease) in the scale parameter ς causes the tails of the distributions to become more distinct. Values in a neighborhood of $\varsigma \simeq 0.50$ marked the global minimum throughout the study.

2. Case 2: Distribution Parameters Estimated.

The Anderson-Darling, Kolmogorov, and Cramér-von Mises goodness-of-fit tests were developed for use in case 1 where distribution parameters are specified, and so precludes their use in the more likely situation where parameters must be estimated. In practice, these procedures are sometimes used anyway with the caveat that the tests are likely to be conservative. Stephens [1974] provides adjustments to the test statistics that enables the tests to be used to test the assumption H_o: $F(x) = F_o(x)$, where $F_o(x) \sim N(\mu, \sigma^2)$ and the population parameters are estimated from the data.

The results of the power study for case 2, are summarized in Figures 16 – 27. As in case 1, the sample sizes are n = 5, 10, 15, and 20, and the level of significance is 0.05. Both location and scale parameters coincide; the scale parameter are values from 0.025 to 3.000 in increments of 0.025.

The power plots are horizontal, demonstrating that power does not change with scale parameter and provides empirical support for Stephens' transformations. Power increases with increasing sample size, as would be expected. When n = 5, the Anderson-Darling test was the least powerful of all three distributions examined. However, none of the distributions had power above 0.30 for this sample size. At the larger sample sizes, the Anderson-Darling test was only slightly better than the Kolmogorov and the Cramér-von Mises tests. In general, when both location and scale parameters agree, all three tests are competitive for the sample sizes and alternative distributions chosen for this study.

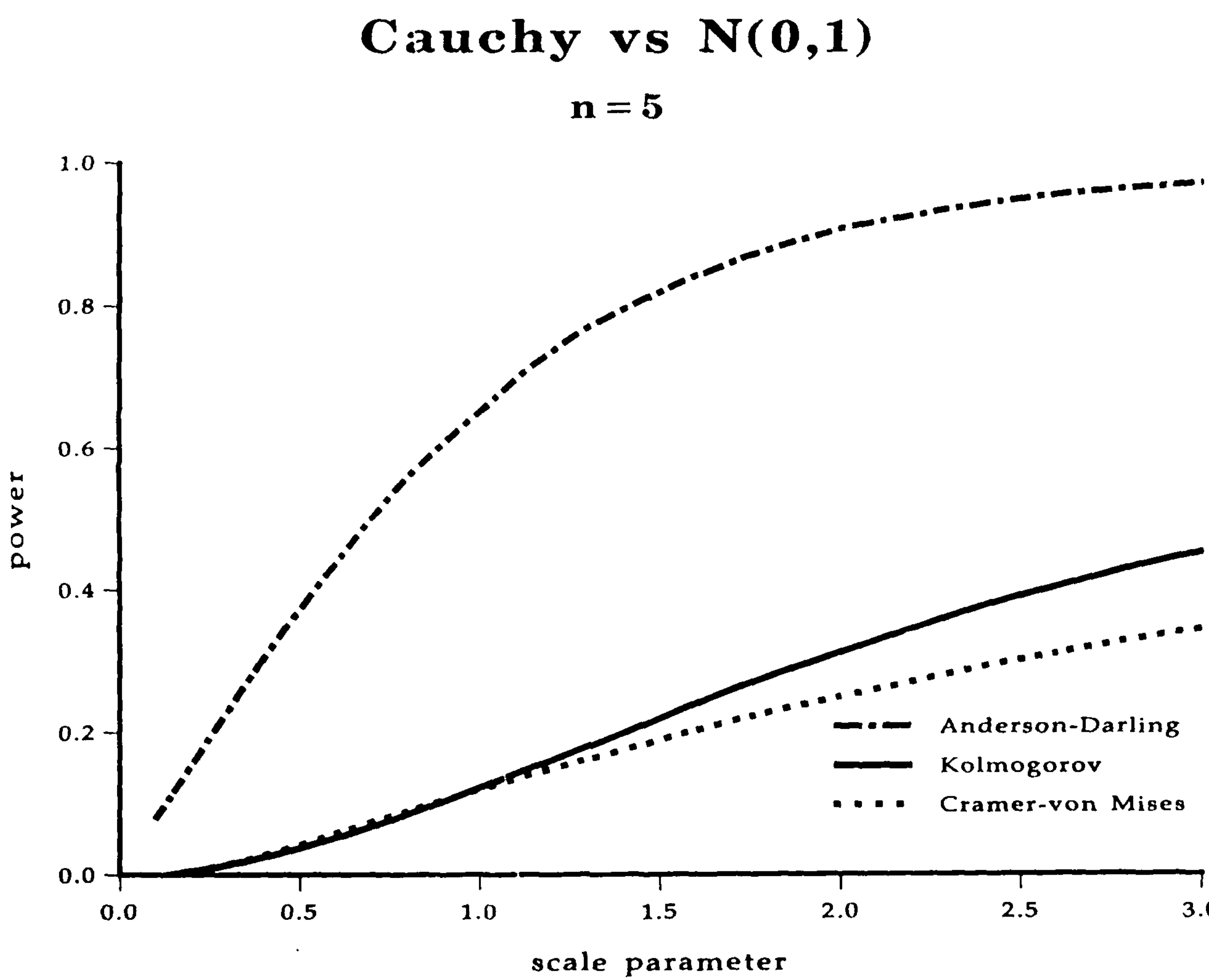

FIGURE 1. Power Plots for Cauchy($0,\varsigma$) vs N(0, 1), n = 5

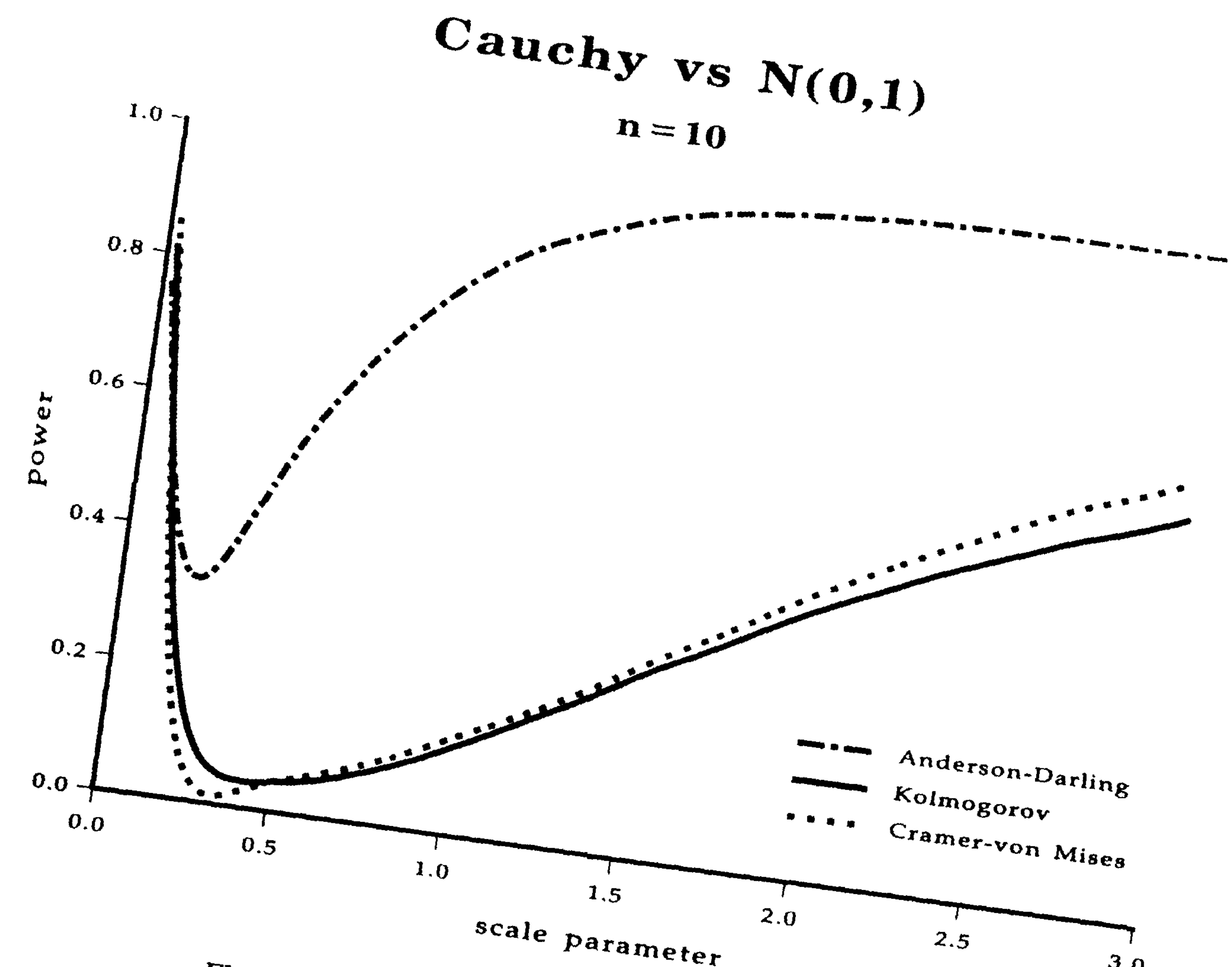

FIGURE 2. Power Plots for Cauchy(0,ς) vs N(0,1), n = 10

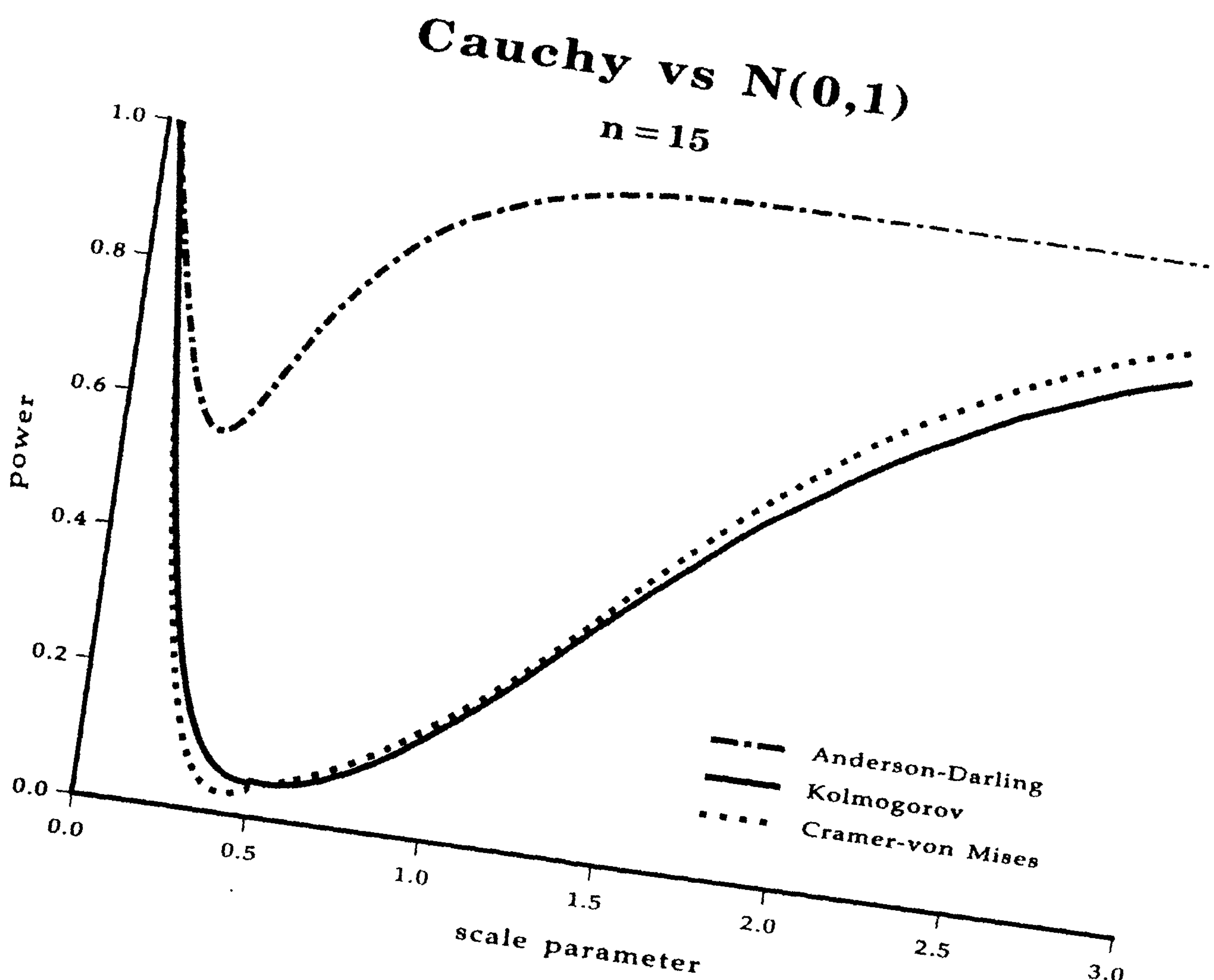

FIGURE 3. Power Plots for Cauchy$(0,\varsigma)$ vs $N(0,1)$, $n = 15$

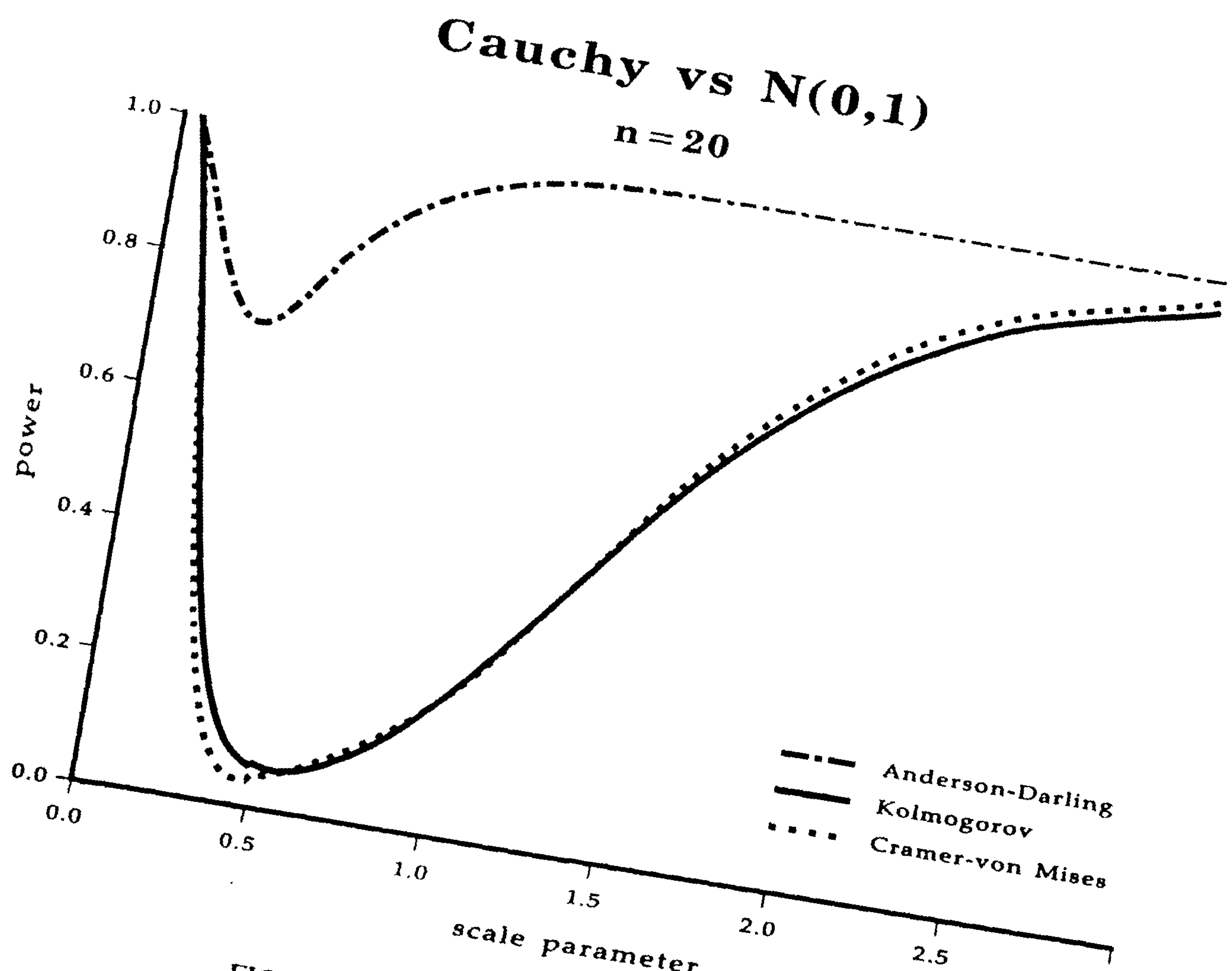

FIGURE 4. Power Plots for Cauchy(0,ς) vs N(0,1), n = 20

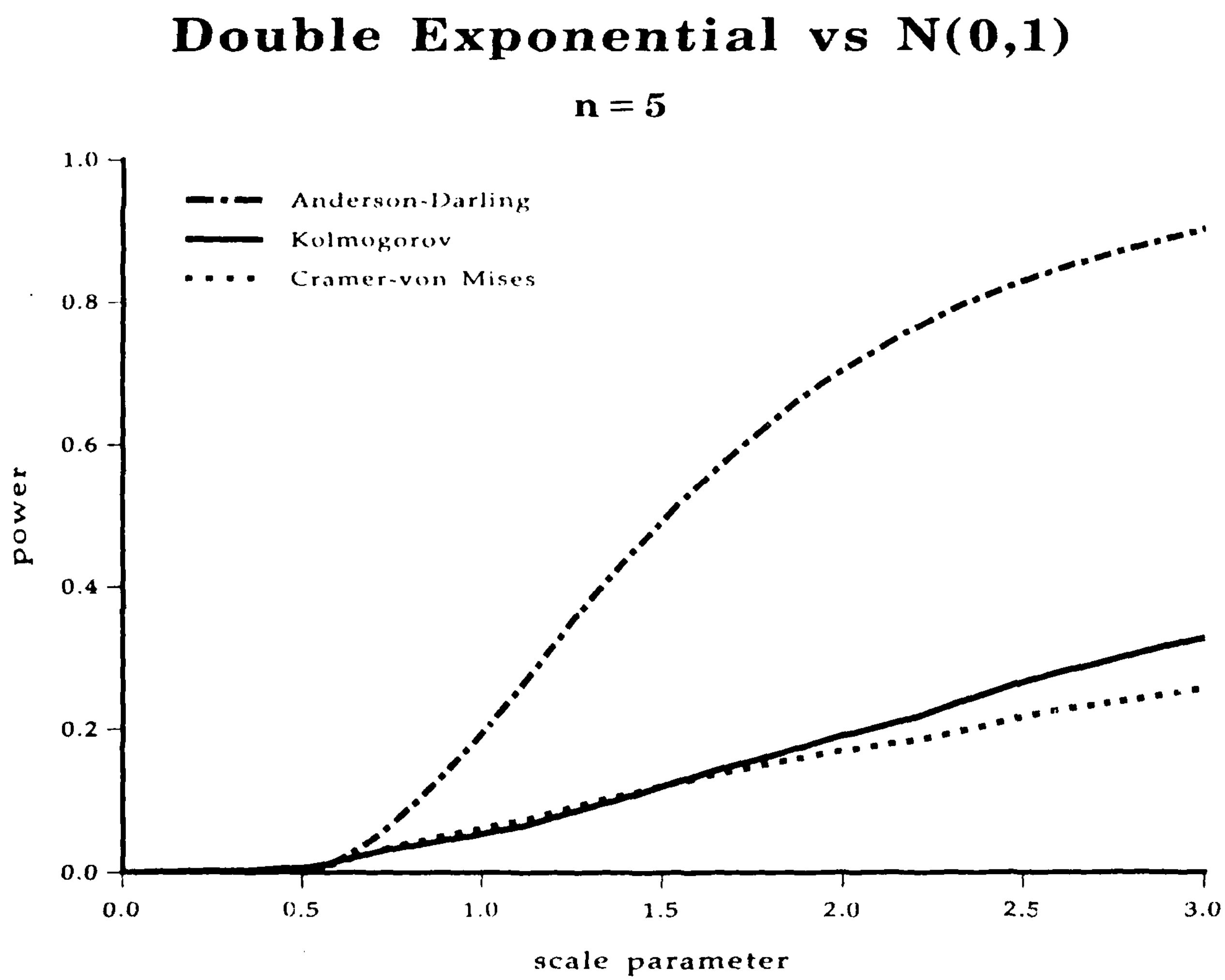

FIGURE 5. Power Plots for Double Exponential(0,ς) vs N(0,1), n = 5

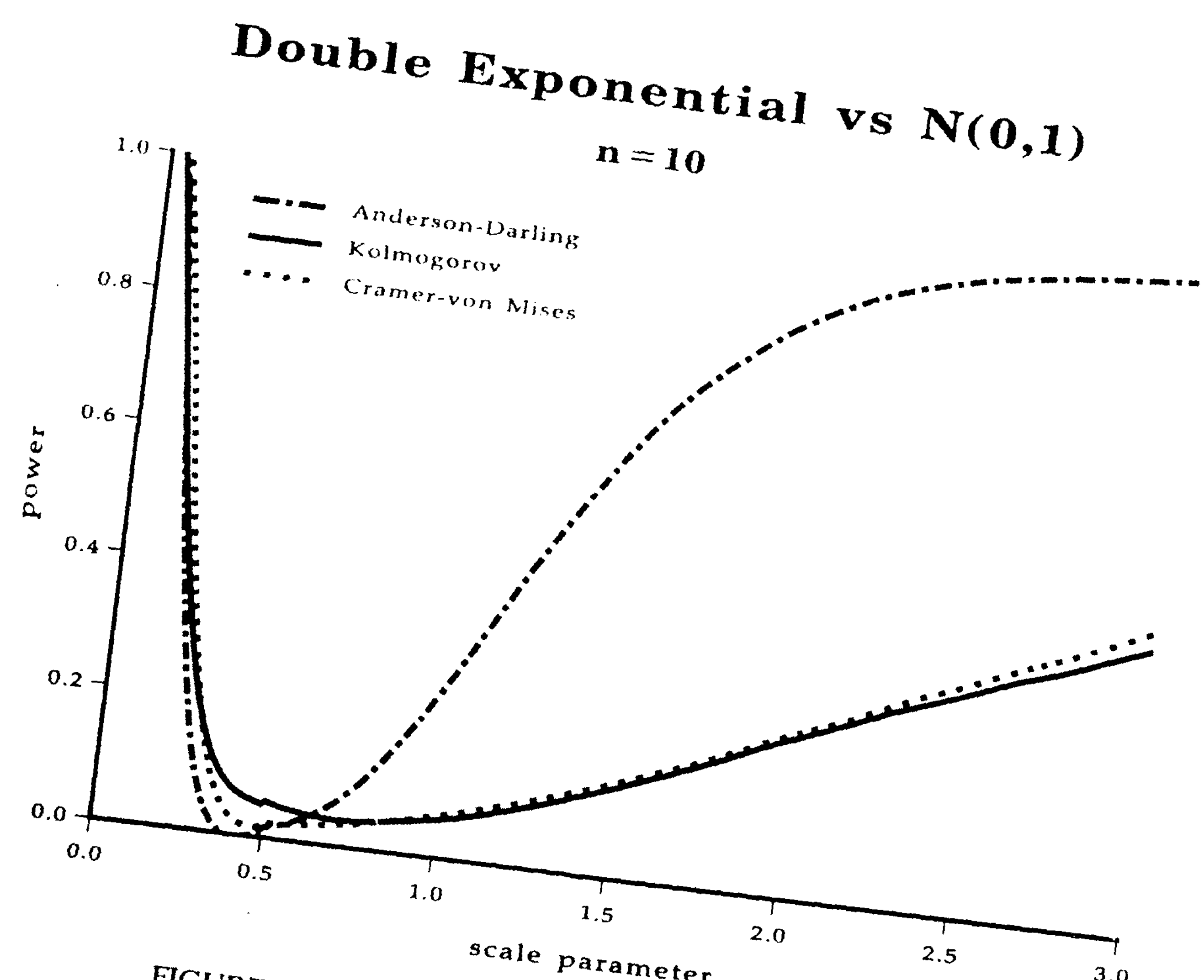

FIGURE 6. Power Plots for Double Exponential(0,ς) vs N(0,1), n = 10

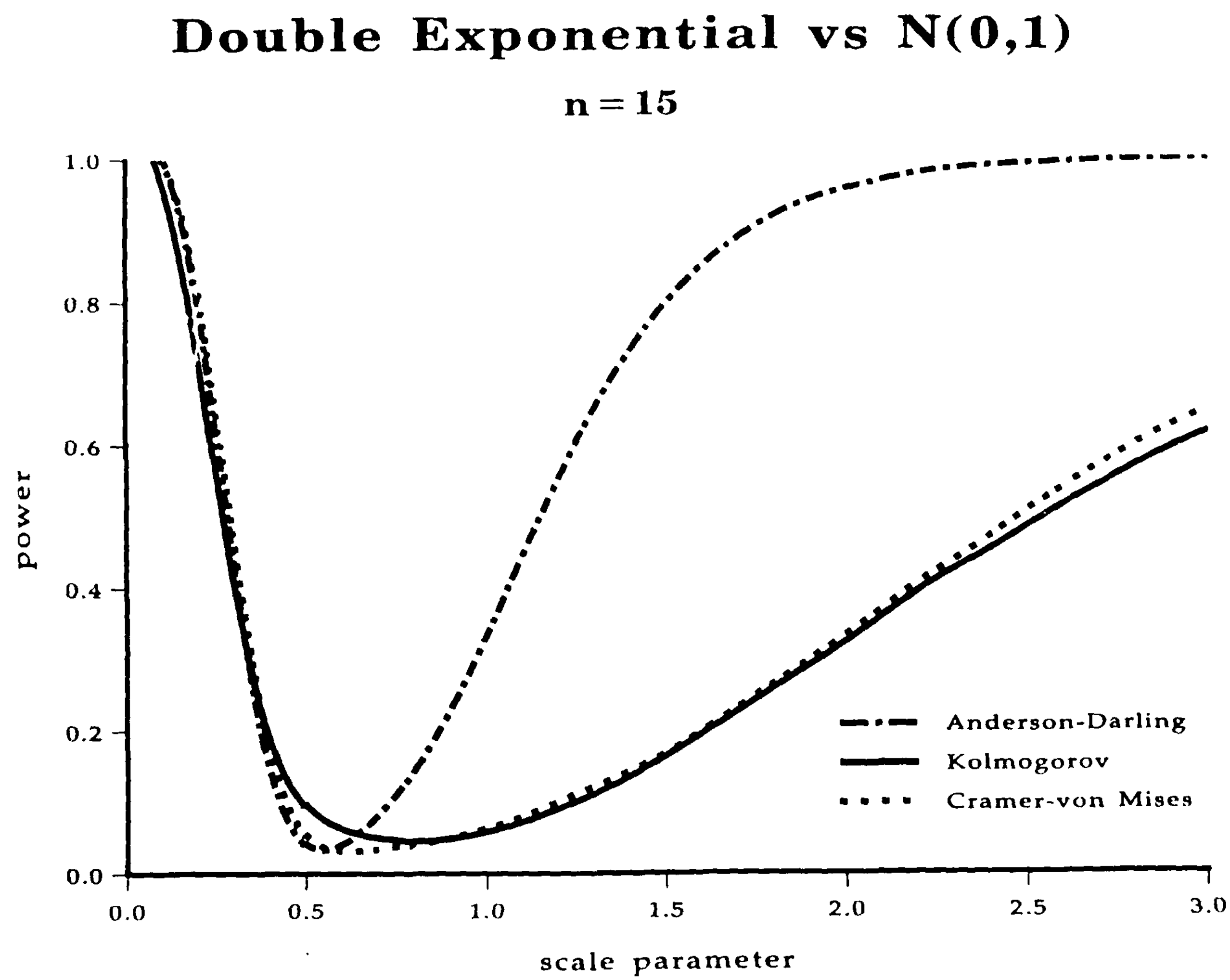

FIGURE 7. Power Plots for Double Exponential(0,ς) vs N(0,1), n = 15

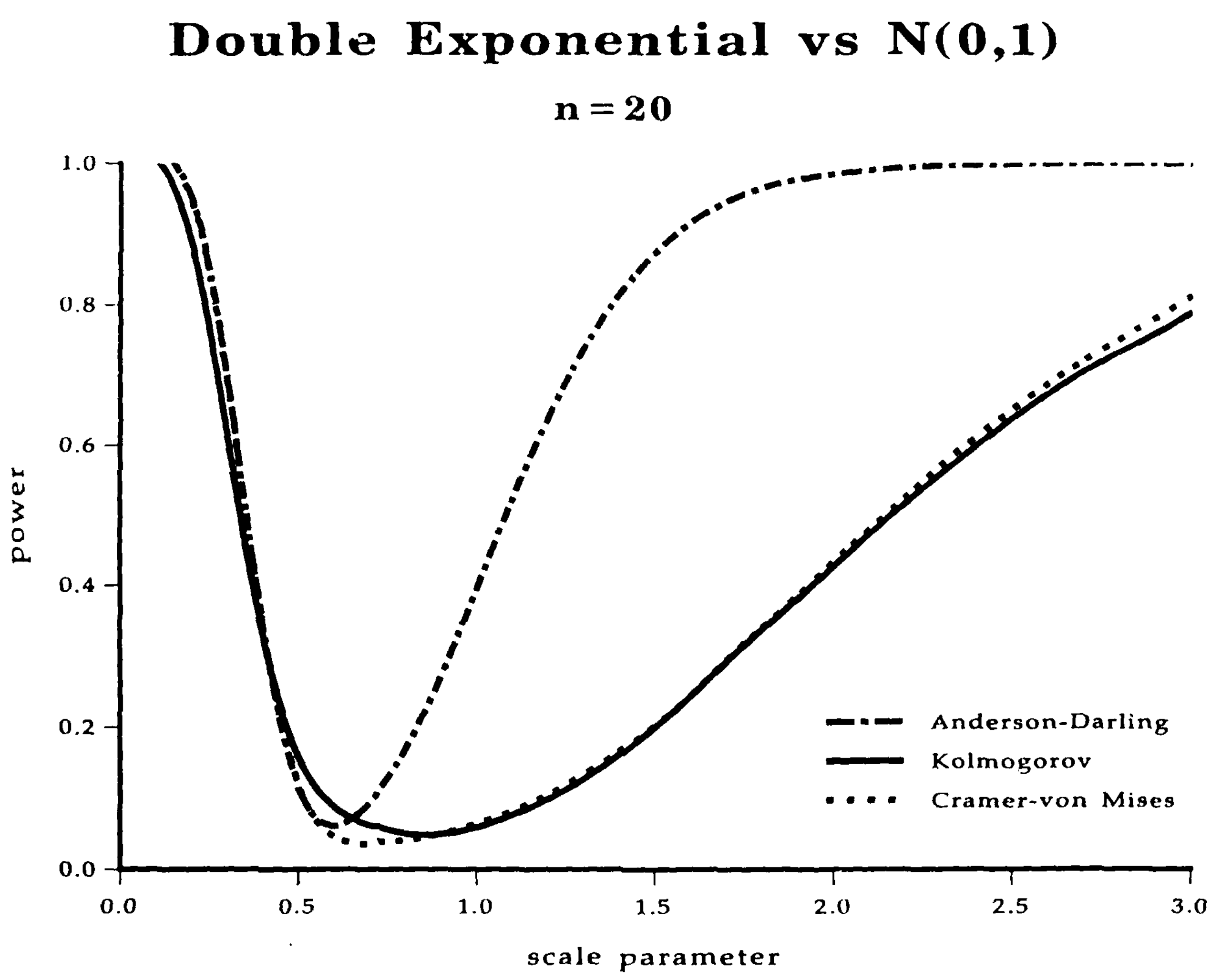

FIGURE 8. Power Plots for Double Exponential(0,ς) vs N(0, 1), n = 20

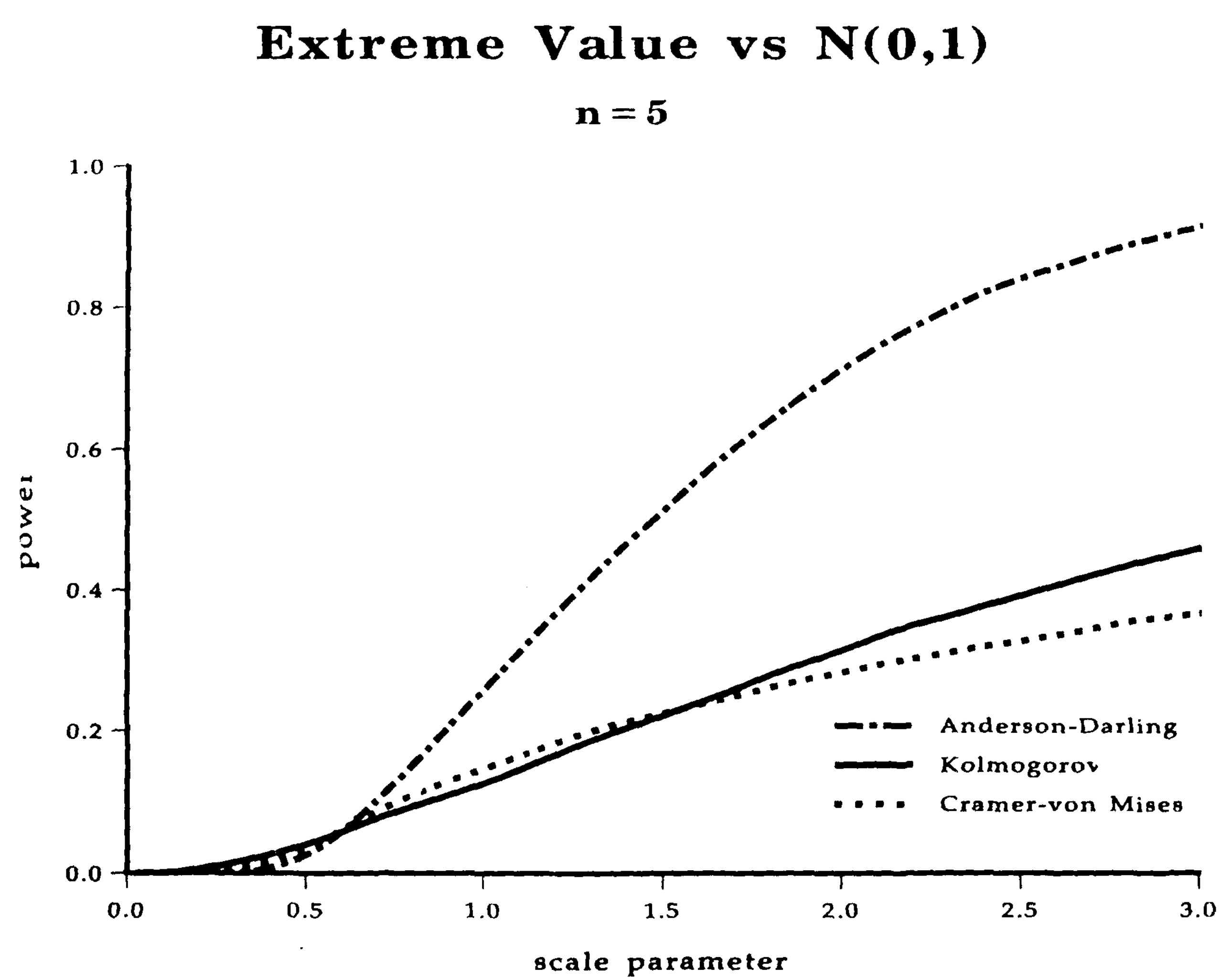

FIGURE 9. Power Plots for Extreme Value($0,\varsigma$) vs N(0, 1), n = 5

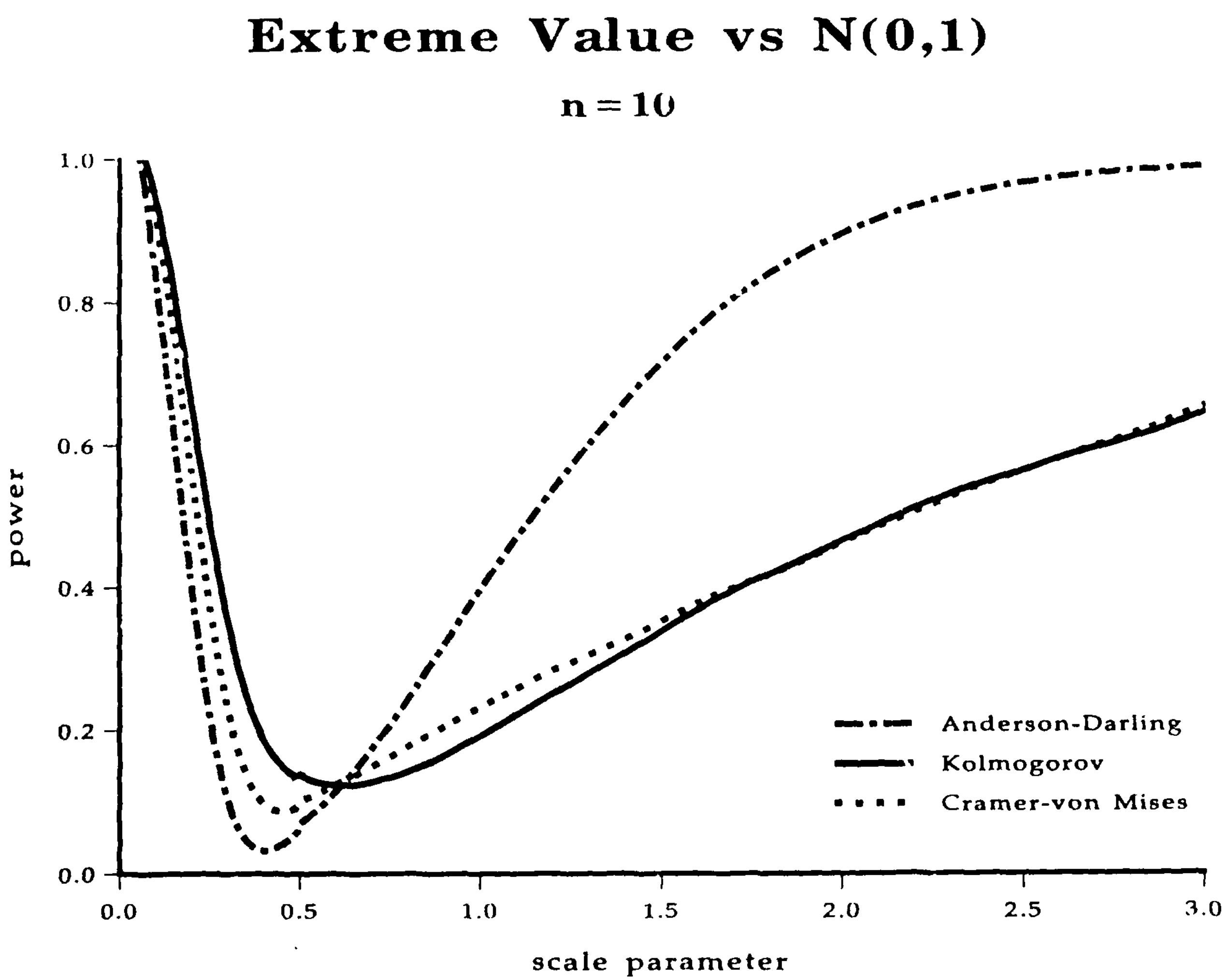

FIGURE 10. Power Plots for Extreme Value(0,ς) vs N(0, 1), n = 10

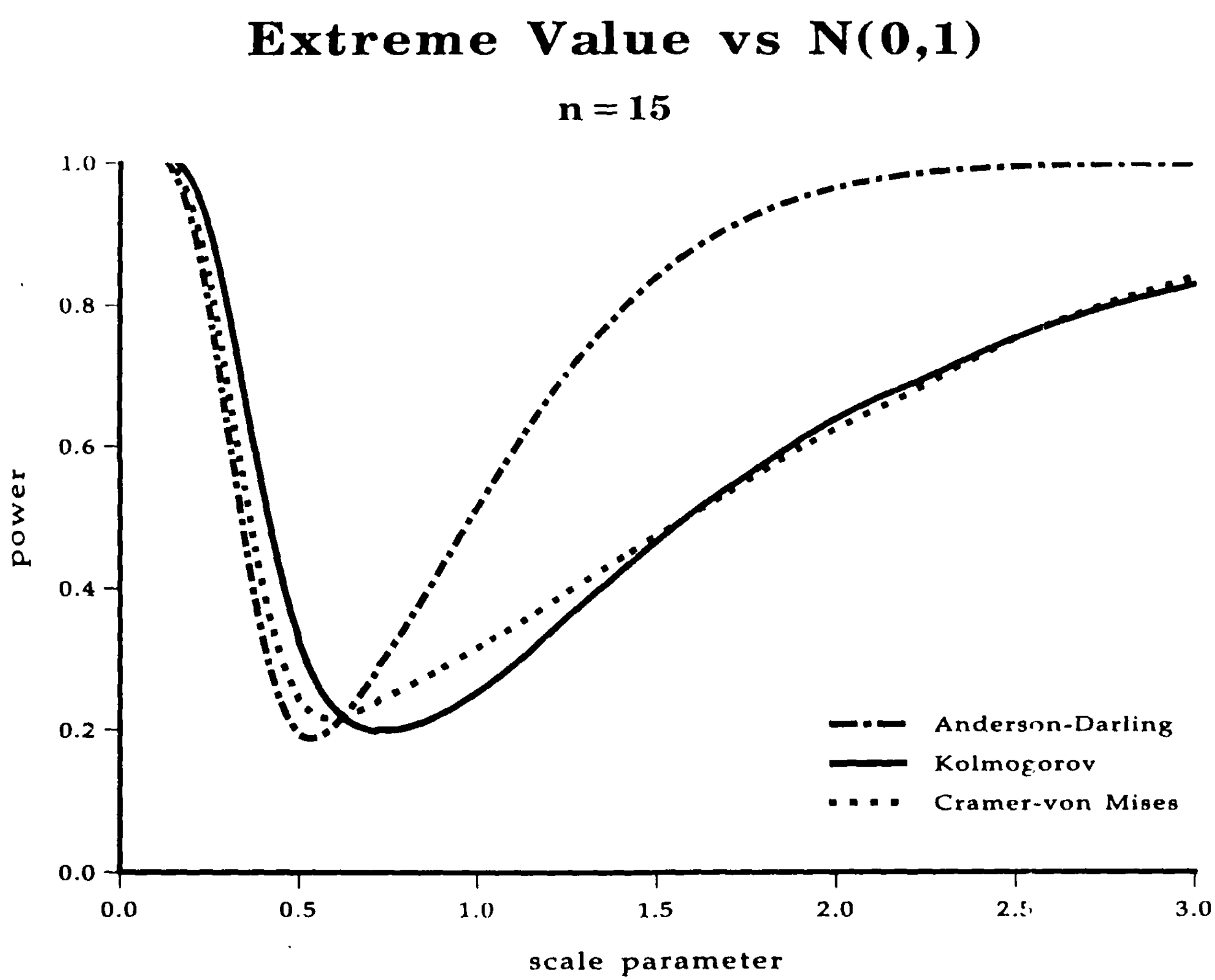

FIGURE 11. Power Plots for Extreme Value($0,\varsigma$) vs N(0, 1), n = 15

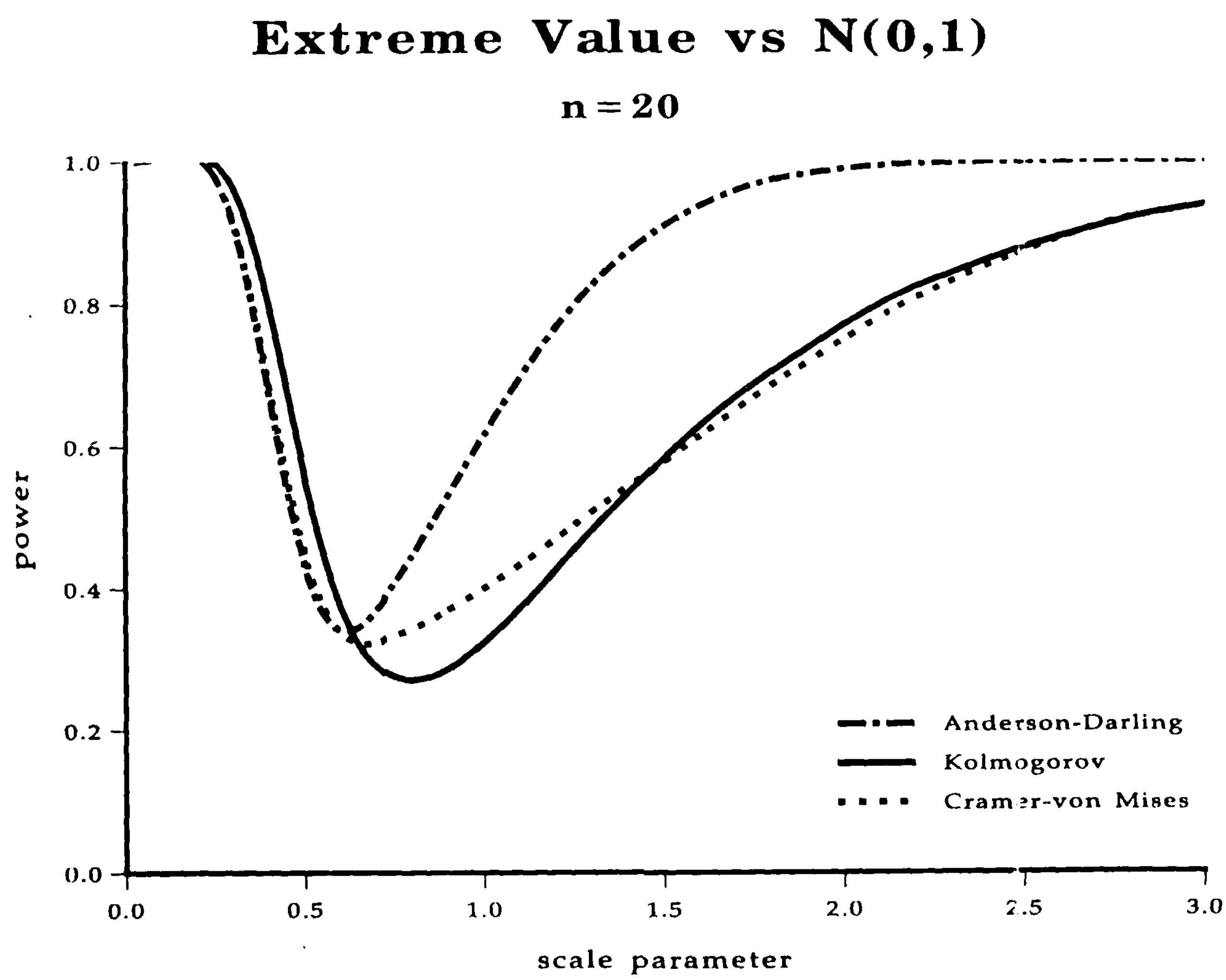

FIGURE 12. Power Plots for Extreme Value(0,ς) vs N(0, 1), n = 20

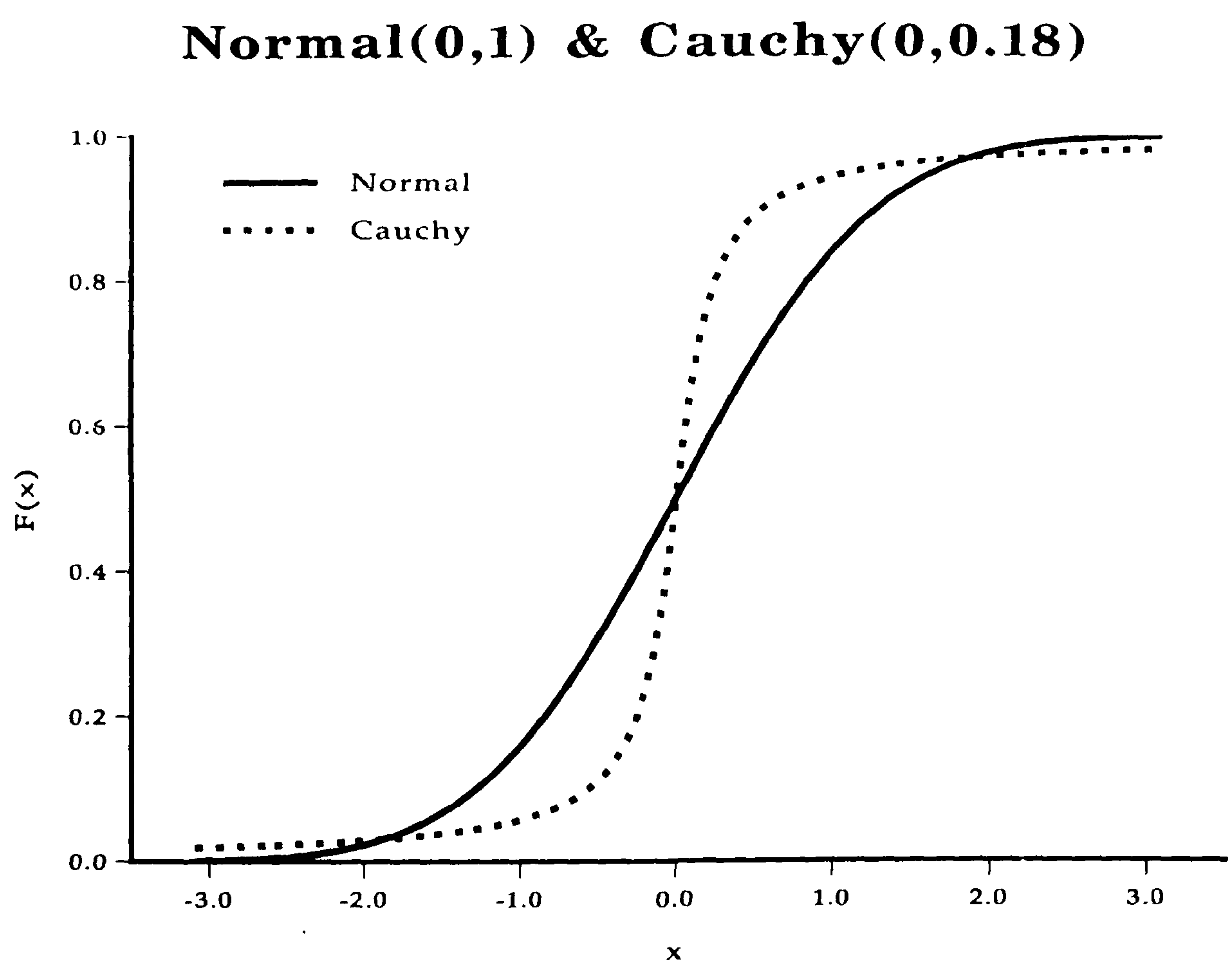

FIGURE 13. Cumulative Distribution Curves, Normal(0,1) & Cauchy(0,0.18)

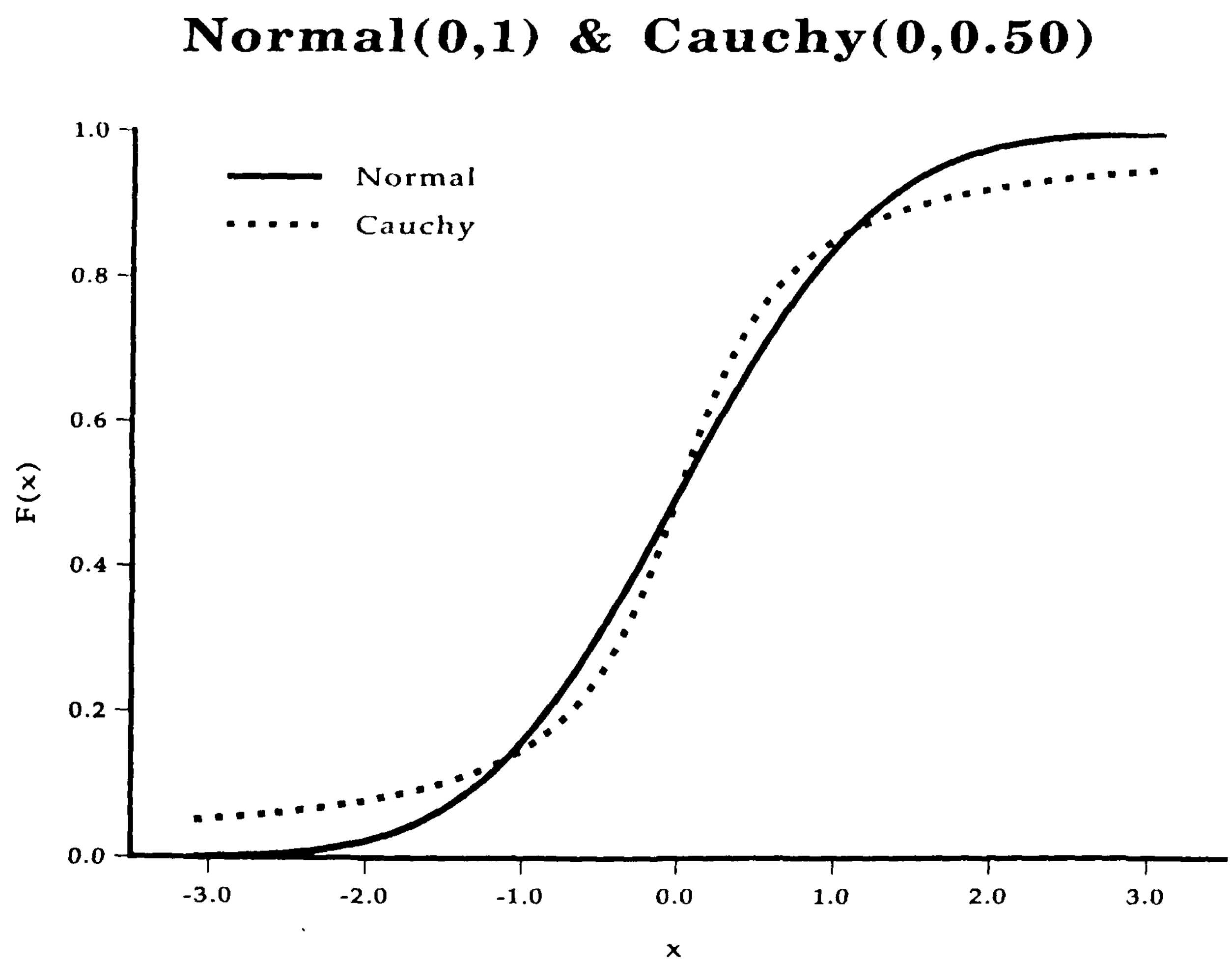

FIGURE 14. Cumulative Distribution Curves, Normal(0,1) & Cauchy(0,0.50)

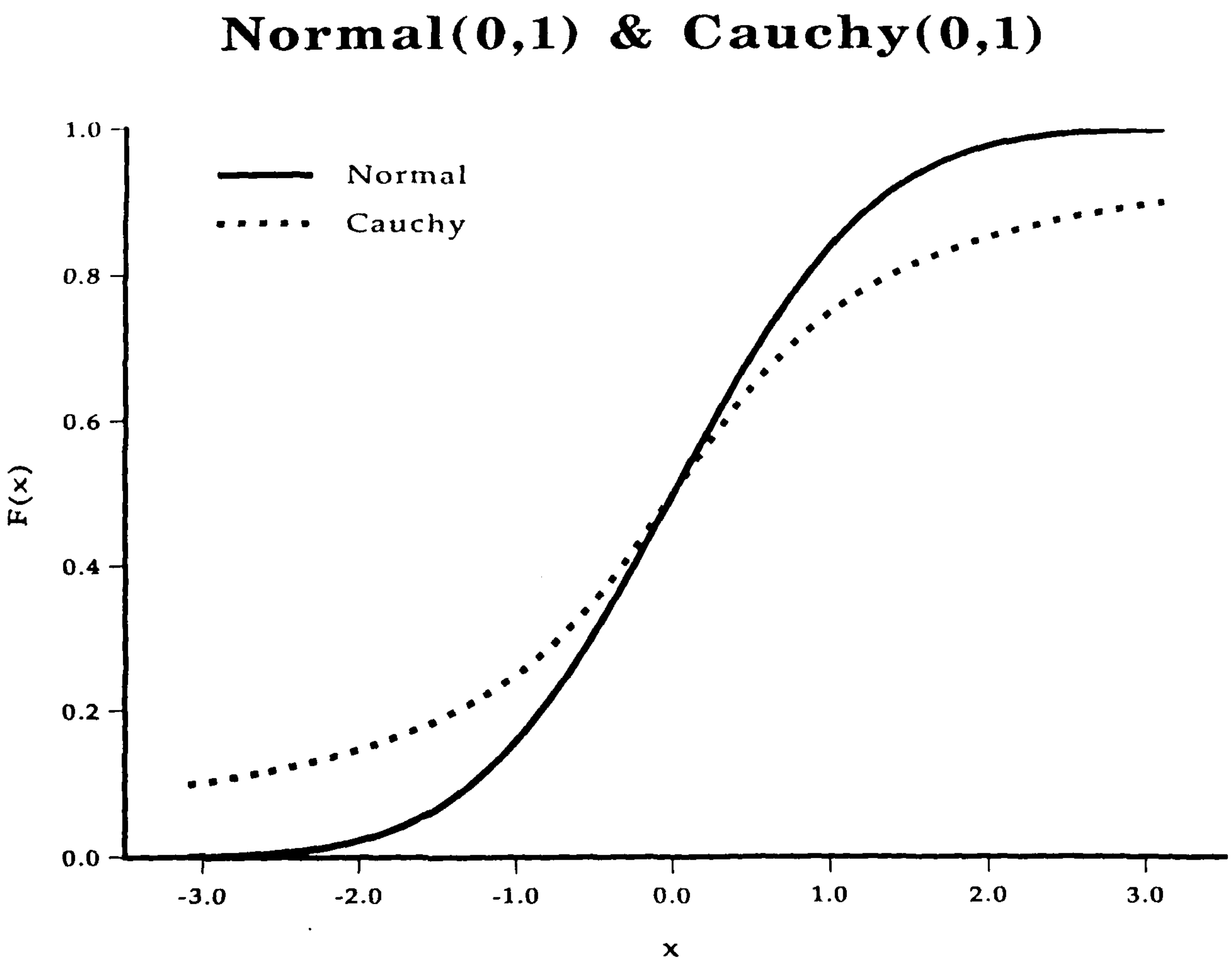

FIGURE 15. Cumulative Distribution Curves, Normal(0,1) & Cauchy(0,1)

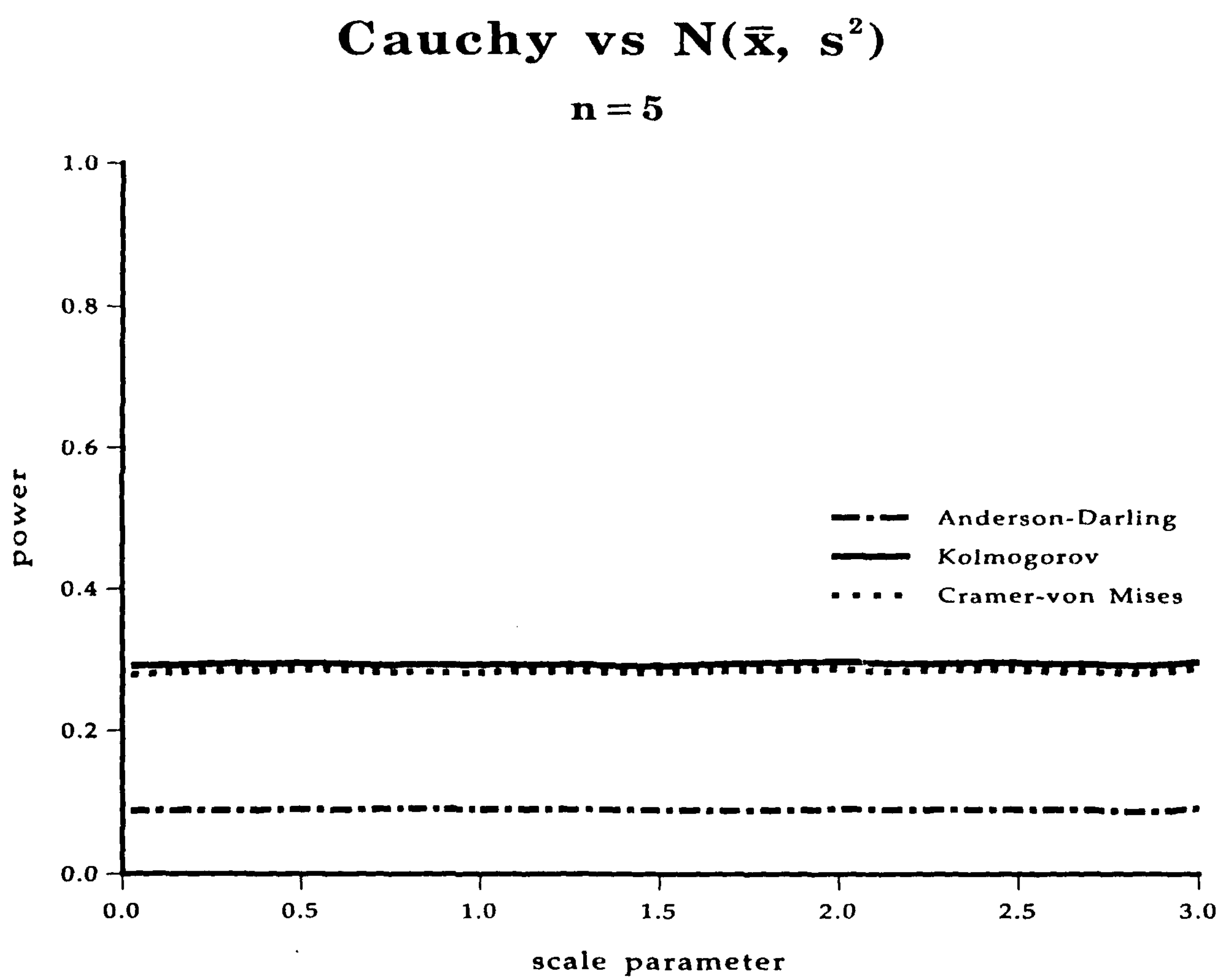

FIGURE 16. Power Plots for Cauchy($0,\varsigma$) vs N($\bar{x}, s^2$), n = 5

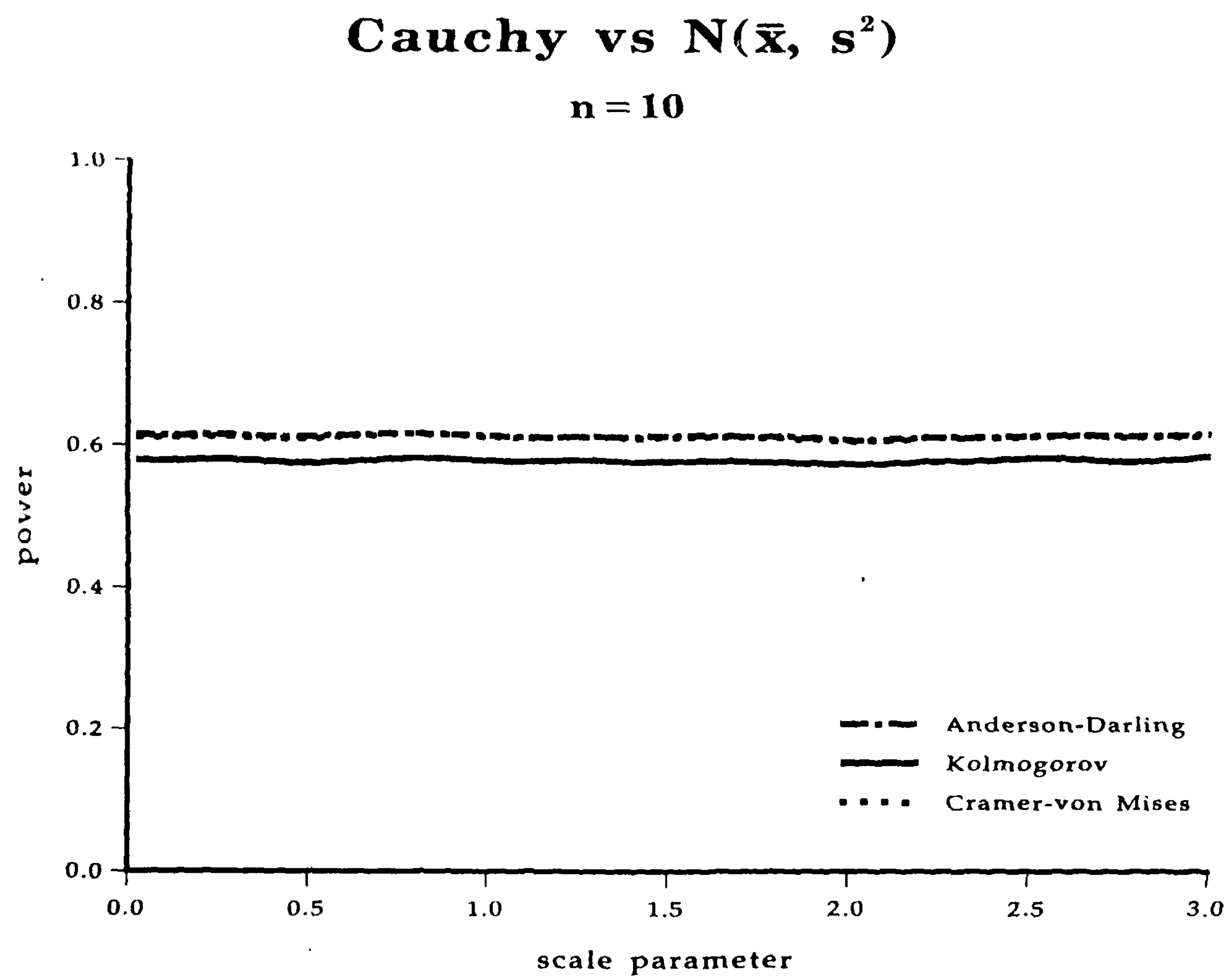

FIGURE 17. Power Plots for Cauchy($0,\varsigma$) vs N($\bar{x},s^2$), n = 10

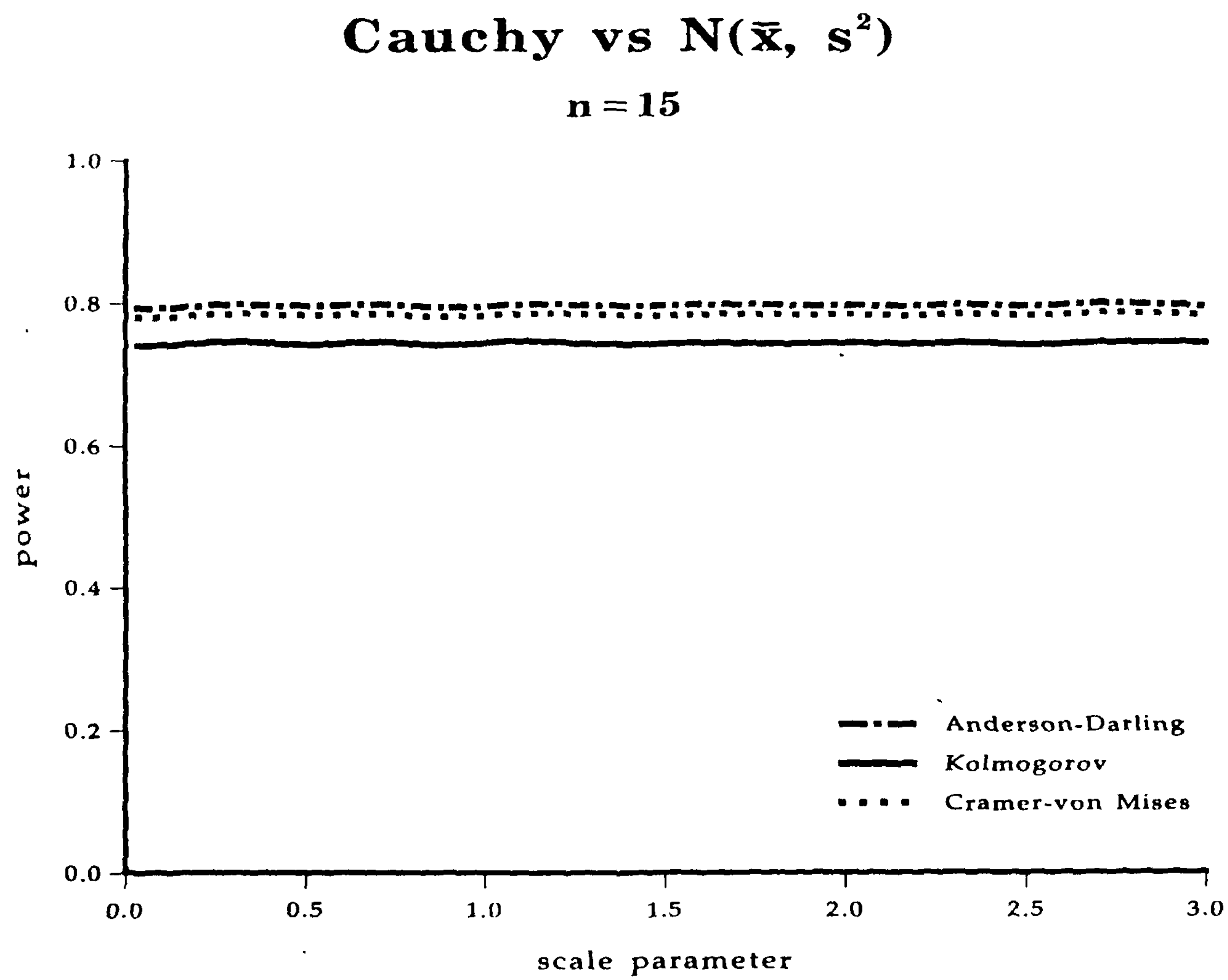

FIGURE 18. Power Plots for Cauchy$(0,\varsigma)$ vs $N(\bar{x}, s^2)$, $n = 15$

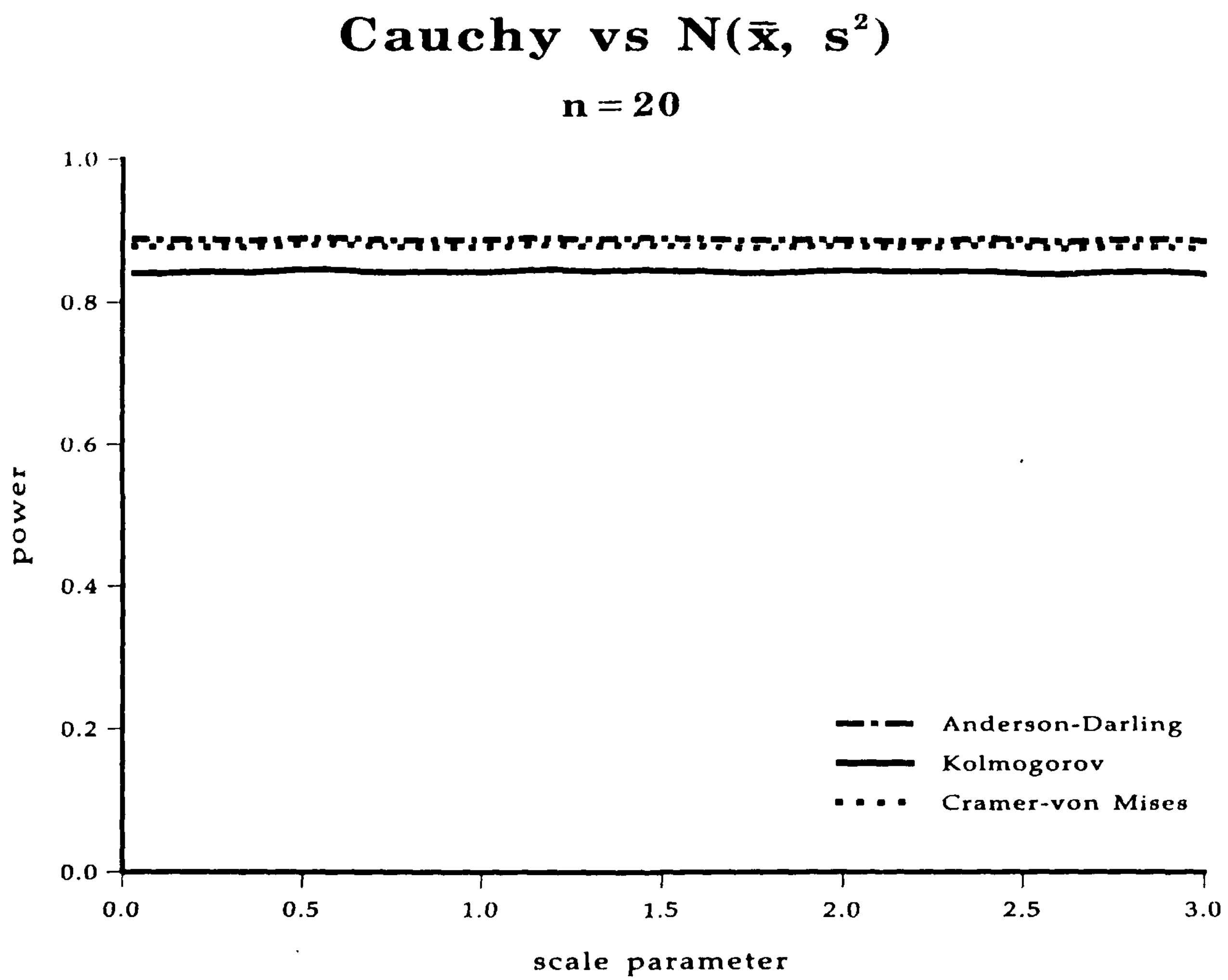

FIGURE 19. Power Plots for Cauchy(0,ς) vs N($\bar{x}$,s^2), n = 20

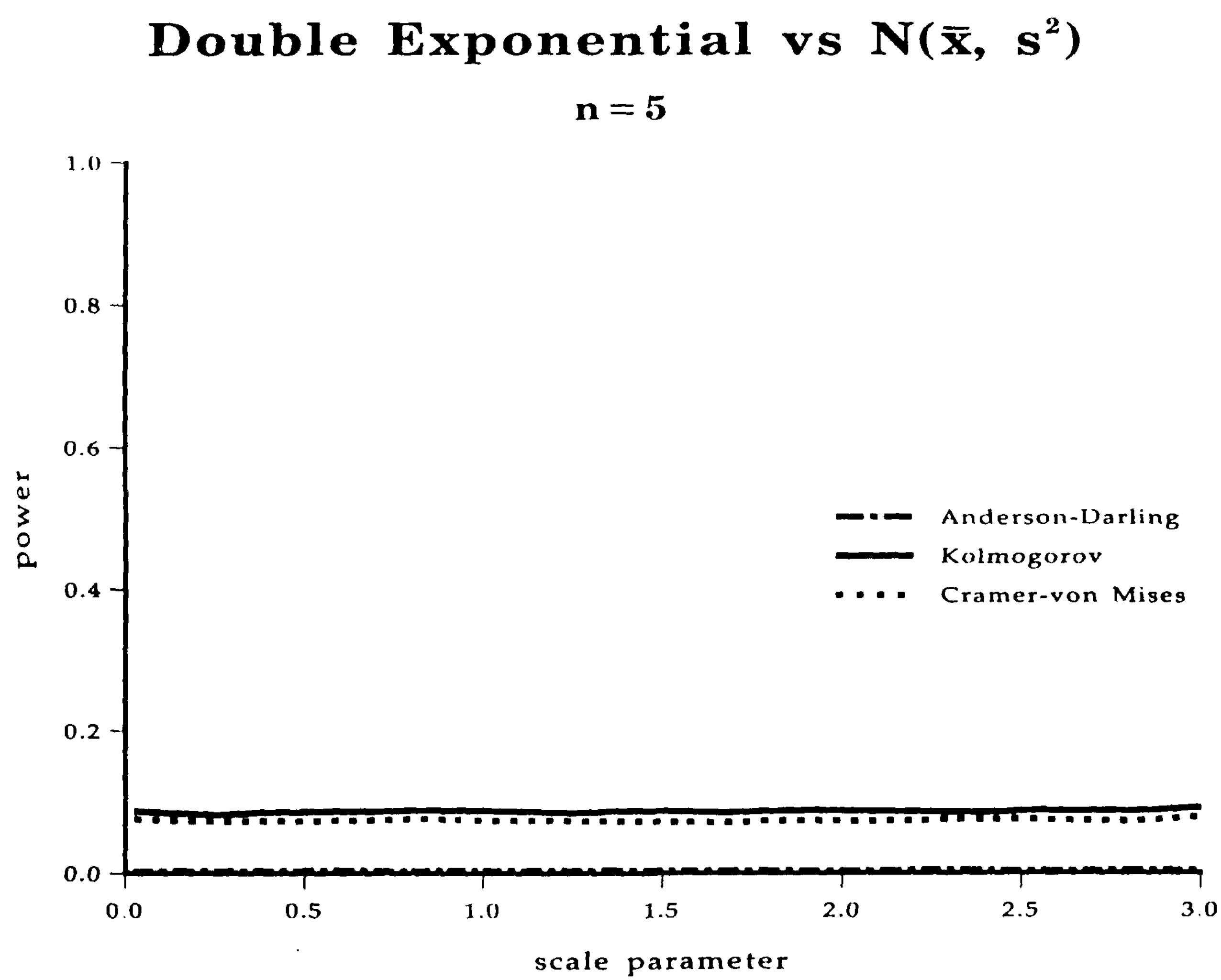

FIGURE 20. Power Plots for Double Exponential(0,ς) vs N($\bar{x}$,s^2), n = 5

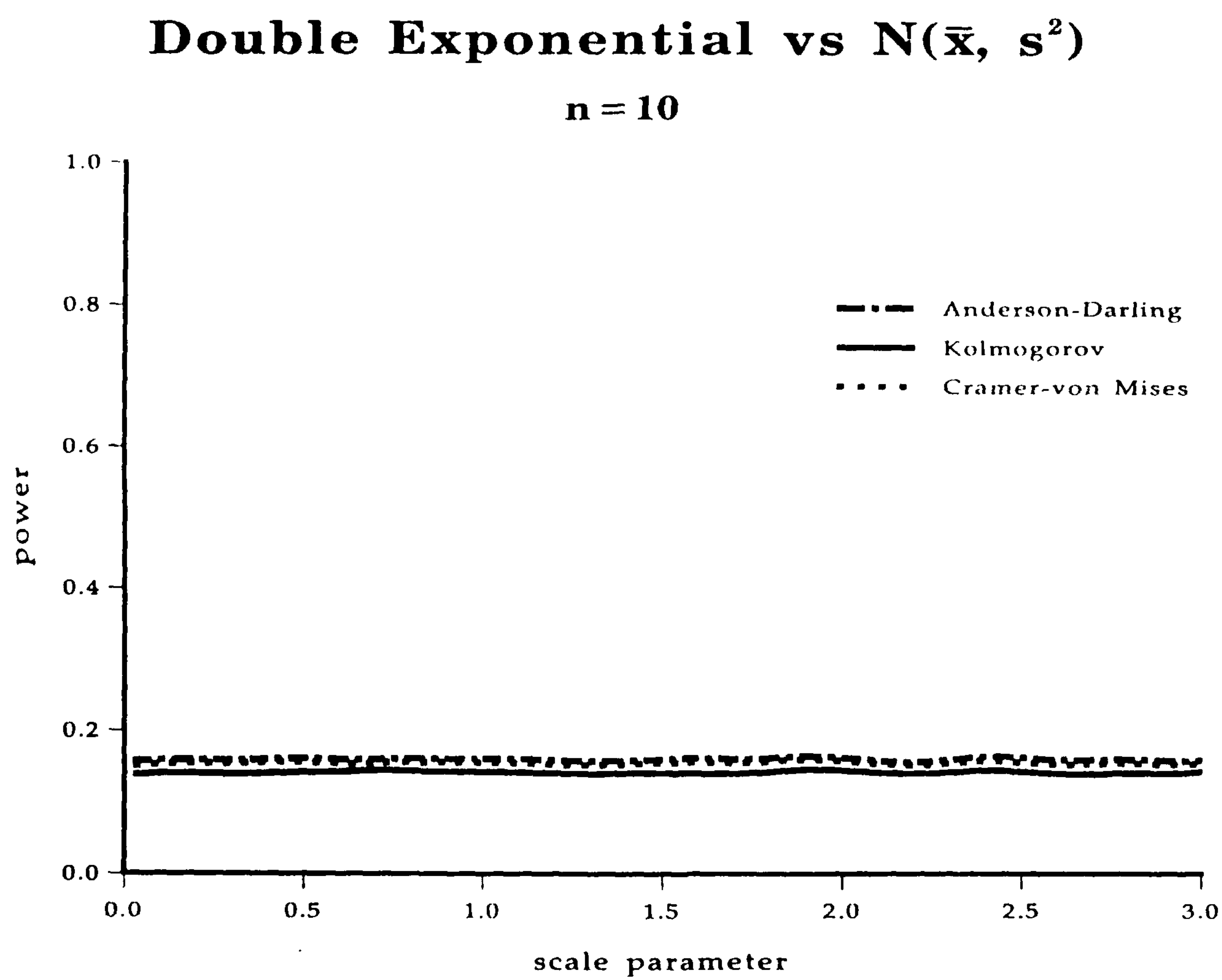

FIGURE 21. Power Plots for Double Exponential(0,ς) vs N($\bar{x}$,s^2), n = 10

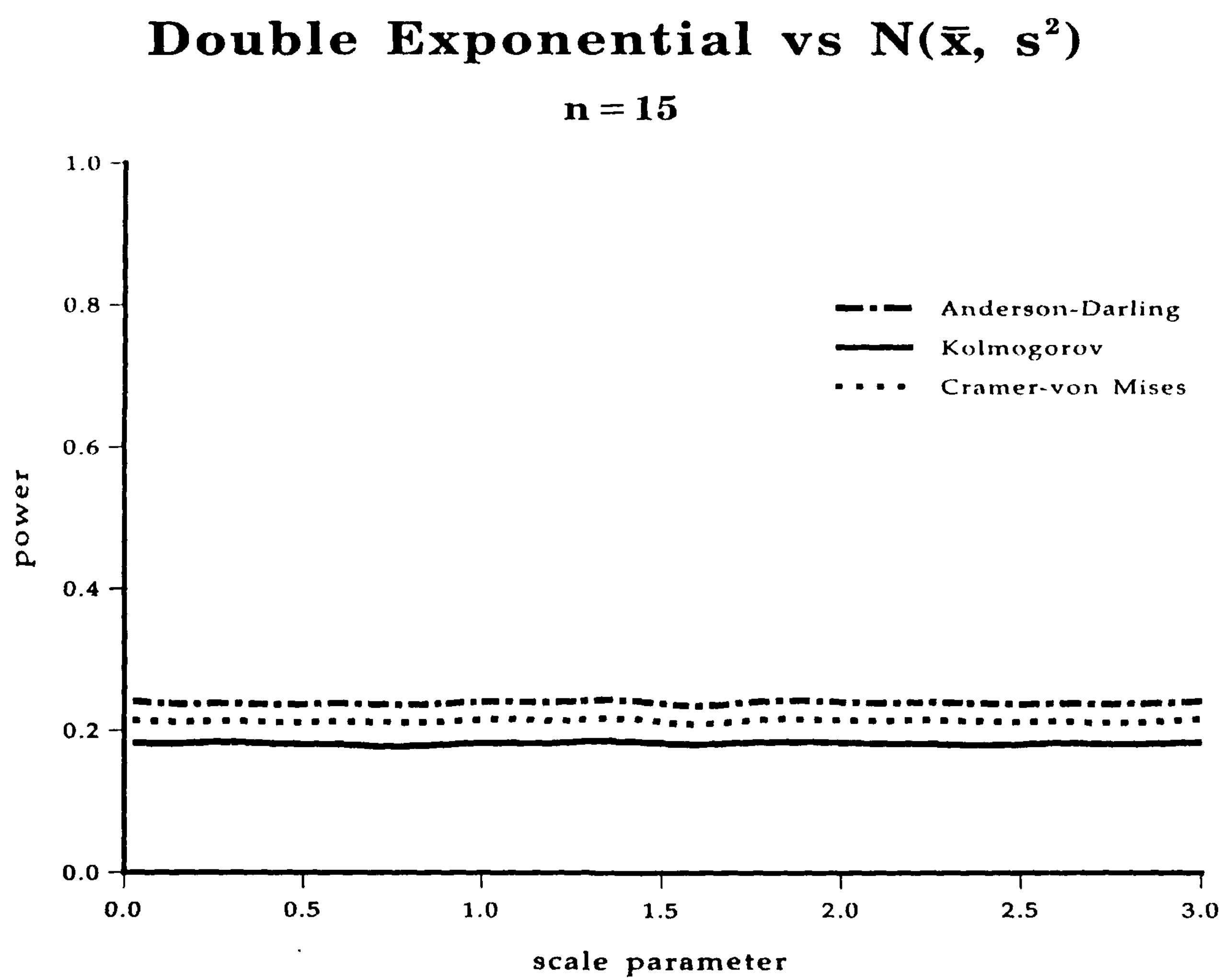

FIGURE 22. Power Plots for Double Exponential(0,ς) vs N($\bar{x}$,s^2), n = 15

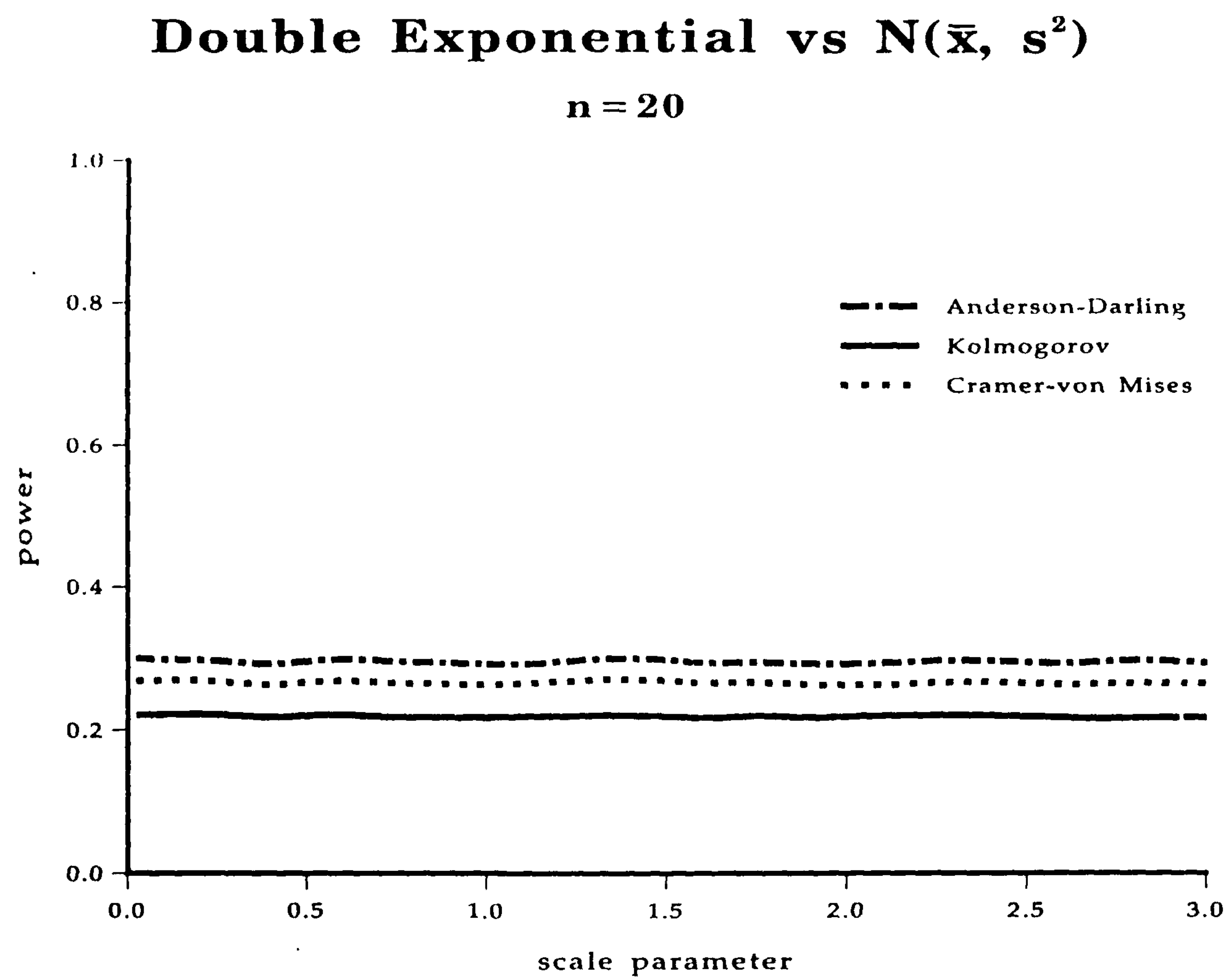

FIGURE 23. Power Plots for Double Exponential$(0,\varsigma)$ vs $N(\bar{x}, s^2)$, $n = 20$

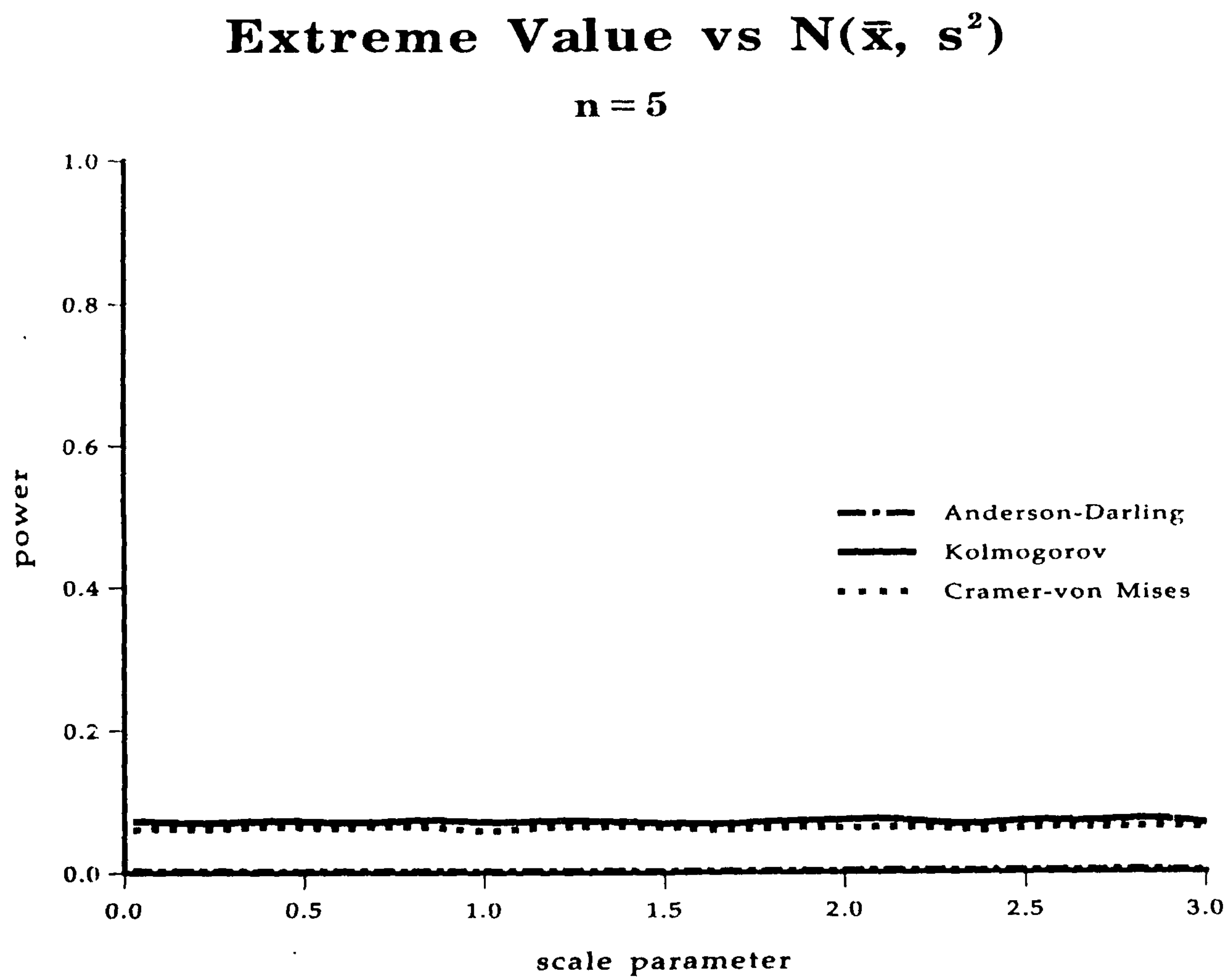

FIGURE 24. Power Plots for Extreme Value($0,\varsigma$) vs N($\bar{x},s^2$), n = 5

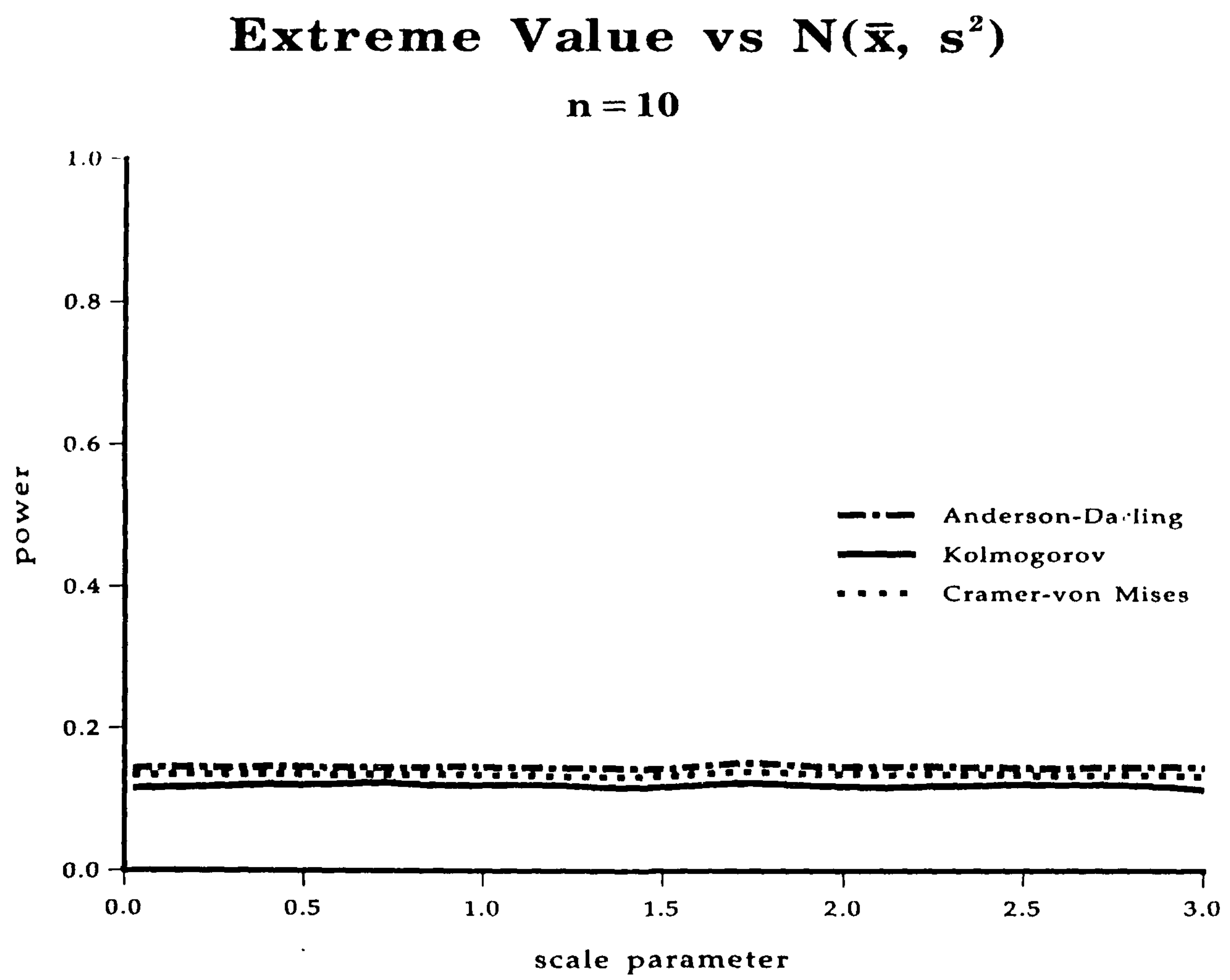

FIGURE 25. Power Plots for Extreme Value$(0,\varsigma)$ vs $N(\bar{x},s^2)$, n = 10

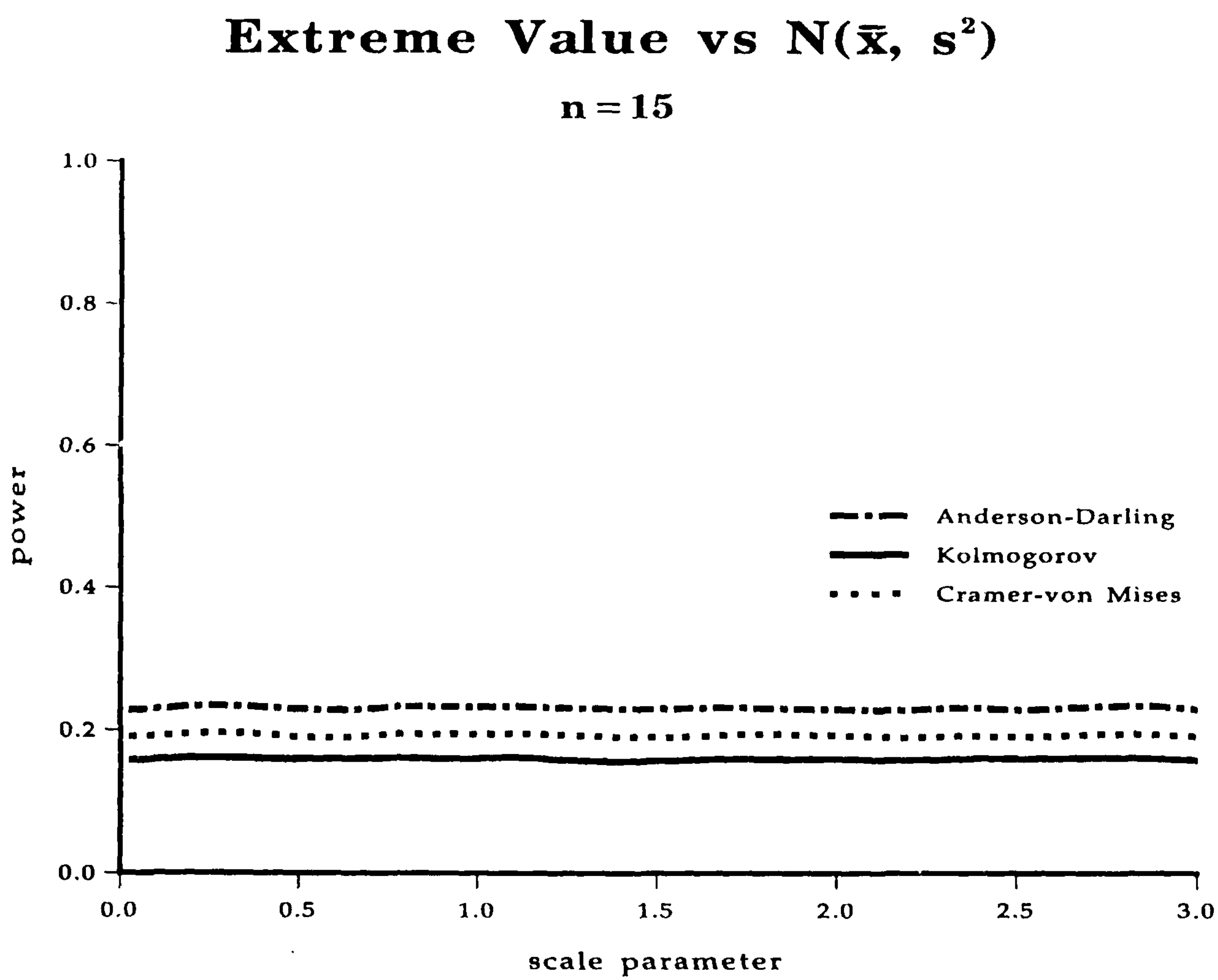

FIGURE 26. Power Plots for Extreme Value$(0,\varsigma)$ vs $N(\bar{x},s^2)$, n = 15

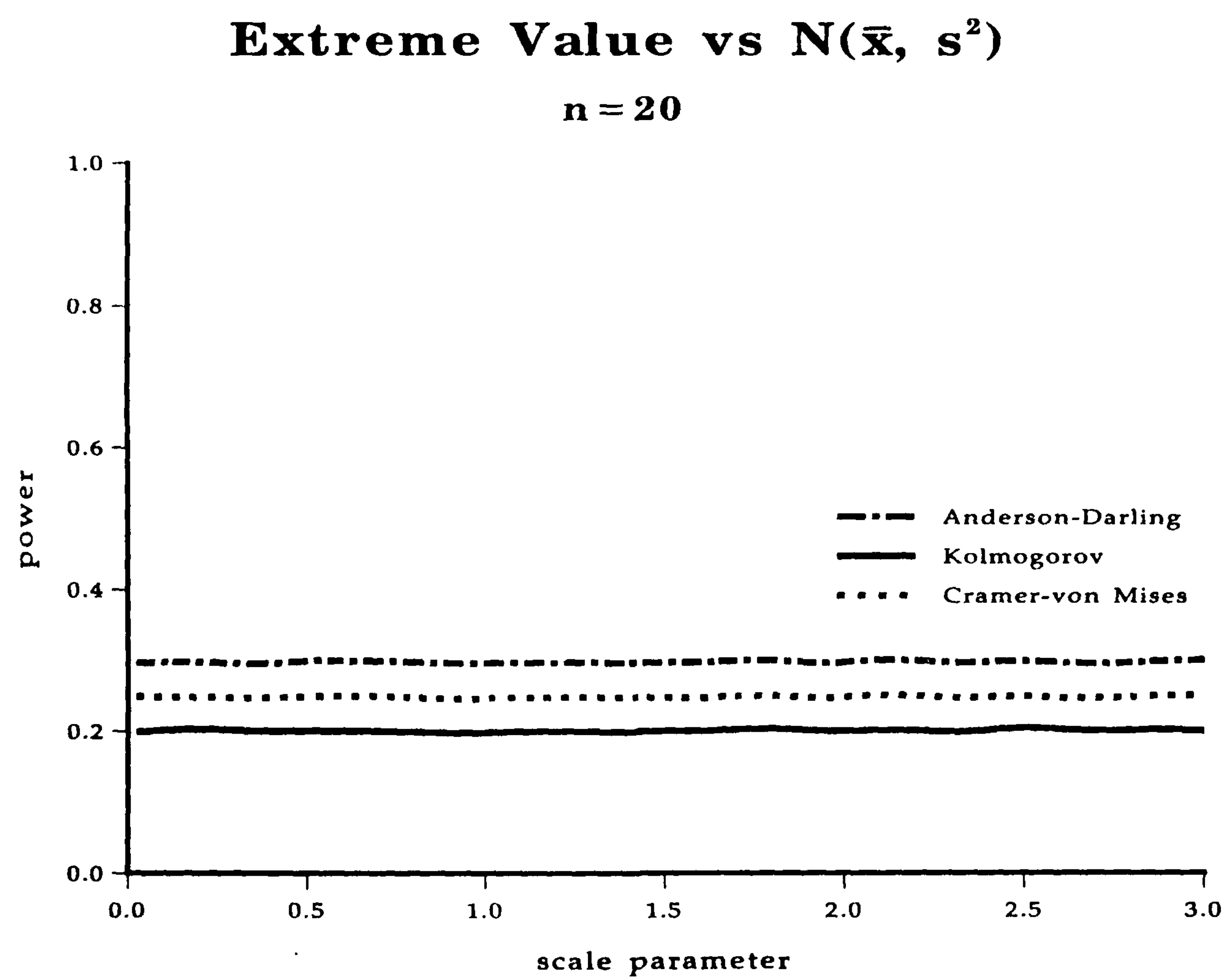

FIGURE 27. Power Plots for Extreme Value$(0,\varsigma)$ vs $N(\bar{x}, s^2)$, $n = 20$

REFERENCES

Abramowitz, Milton and Irene A. Stegun, eds. *Handbook of Mathematical Functions with Formulas, Graphs, and Mathematical Tables.* National Bureau of Standards, Applied Mathematics Series 55, June 1964.

Anderson, T. W. and D. A. Darling. "Asymptotic Theory of Certain 'Goodness of Fit' Criteria Based on Stochastic Process." *Annals of Mathematical Statistics,* 1952, Vol. 23, pp. 193-212.

----------. "A Test of Goodness of Fit." *Journal of the American Statistical Association,* 1954, Vol. 49, pp. 765-769.

Cogan, Edward J. and Robert Z. Norman. *Handbook of Calculus, Difference and Differential Equations.* Englewood Cliffs, N.J.: Prentice-Hall, 1958.

Conover, W. J. *Practical Nonparametric Statistics.* Second Edition. New York: John Wiley & Sons, 1980.

Cramér, Harald. "On the composition of elementary errors." *Skandinavisk Aktuarietidskrift,* 1928, Vol. 11, pp. 13-74, 141-180.

Feller, William. *An Introduction to Probability Theory and its Application, Volume II.* New York: John Wiley and Sons, 1966.

Harter, H. Leon. "Modified Asymptotic Formulas for Critical Values of the Kolmogorov Test Statistic." *The American Statistician,* May 1980, Vol. 34, No. 2, pp. 110-111.

Johnson, Norman L. and Samuel Kotz. *Continuous Univariate Distributions - 1. Distributions in Statistics.* New York: John Wiley & Sons, 1970.

Jolley, L. B. W. *Summation of Series.* 2nd Revised Edition. New York: Dover Publications, Inc., 1961.

Kernighan, Brain W. and Dennis M. Ritchie. *The C Programming Language.* New Jersey: Prentice-Hall, 1978.

Kolmogorov, A. N. "Sulla Determinazione Empirica di una Leggi di Distributione." *Giornale dell' Istituto Italiano degli Attuari,* 1933, Vol. 4, pp. 83-91.

Lewis, Peter A. W. "Distribution of the Anderson-Darling Statistic." *The Annals of Mathematical Statistics,* 1961, Vol. 32, pp. 1118-1124.

Miller, R. G., Jr. "Table of Percentage Points of Kolmogorov Statistics." *Journal of the American Statistical Association*, 1956, Vol. 51, pp. 111-112.

Rubinstein, Reuven Y. *Simulation and the Monte Carlo Method.* New York: John Wiley & Sons, 1981.

Stephens, M. A. "EDF Statistics for Goodness of Fit and Some Comparisons." *Journal of the American Statistical Association*, September 1974, Vol. 69, No. 347, pp. 730-737.

----------. "Use of the Kolmogorov-Smirov, Cramér-Von Mises and Related Statistics without Extensive Tables." *Royal Statistical Society, Series B*, 1970, Vol. 32, No. 1, pp. 115-122.

---------- and Urs R. Maag. "Further Percentage Points for W_N^2." *Biometrika*, 1968, Vol. 55, pp. 428-430.

von Mises, R. *Wahrscheinlichkeitsrechnung und Ihre Anwendung in der Statistik und Theoretischen Physik*. Leipzig: F. Deuticke, 1931.

INTENTIONALLY LEFT BLANK.

APPENDIX A: EXPECTATION OF SQUARED DISCREPANCY BETWEEN AN EMPIRICAL DISTRIBUTION FUNCTION AND A SPECIFIED DISTRIBUTION FUNCTION

INTENTIONALLY LEFT BLANK.

APPENDIX A: EXPECTATION OF SQUARED DISCREPANCY BETWEEN AN EMPIRICAL DISTRIBUTION FUNCTION AND A SPECIFIED DISTRIBUTION FUNCTION

$$E\left[F_n(x) - F(x)\right]^2 = E\left[F_n(x) - H(x) + H(x) - F(x)\right]^2$$

$$= E\left[\left\{F_n(x) - H(x)\right\} - \left\{F(x) - H(x)\right\}\right]^2$$

$$= E\left[\left\{F_n(x) - H(x)\right\}^2 - 2\left\{F_n(x) - H(x)\right\}\left\{F(x) - H(x)\right\} + \left\{F(x) - H(x)\right\}^2\right]$$

$$= E\left[F_n(x) - H(x)\right]^2 - 2\left\{F(x) - H(x)\right\}E\left[F_n(x) - H(x)\right] + \left\{F(x) - H(x)\right\}^2$$

$$= E\left[F_n(x) - H(x)\right]^2 + \left\{F(x) - H(x)\right\}^2$$

$$= E\left[F_n^2(x) - 2\,F_n(x)H(x) + H^2(x)\right] + \left\{F(x) - H(x)\right\}^2$$

$$= E\left[F_n^2(x)\right] - 2\,H(x)\,E\left[F_n(x)\right] + H^2(x) + \left\{F(x) - H(x)\right\}^2$$

$$= \left[\frac{1}{n} H(x)\{1 - H(x)\} + H^2(x) - 2\,H^2(x) + H^2(x) + \left\{F(x) - H(x)\right\}^2\right]$$

$$E\left[F_n(x) - F(x)\right]^2 = \left[\frac{1}{n}\left\{H(x)\{1 - H(x)\}\right\} + \left\{F(x) - H(x)\right\}^2\right].$$

APPENDIX B: EXPANSION AND INTEGRATION OF THE ANDERSON-DARLING STATISTIC

INTENTIONALLY LEFT BLANK.

APPENDIX B: EXPANSION AND INTEGRATION OF THE ANDERSON-DARLING STATISTIC

$$W_n^2 = n \int_{-\infty}^{\infty} \frac{[F_n(x) - F(x)]^2}{F(x)[1 - F(x)]} \, dF(x)$$

$$= n \left\{ \int_{-\infty}^{x_1} \frac{[F_n(x) - F(x)]^2}{F(x)[1 - F(x)]} \, dF(x) + \sum_{k=1}^{n-1} \int_{x_k}^{x_{k+1}} \frac{[F_n(x) - F(x)]^2}{F(x)[1 - F(x)]} \, dF(x) \right.$$

$$\left. + \int_{x_n}^{\infty} \frac{[F_n(x) - F(x)]^2}{F(x)[1 - F(x)]} \, dF(x) \right\}$$

Let $F(x) = u$
$dF(x) = du$
$F(x_i) = u_i$

$$W_n^2 = n \left\{ \int_0^{u_1} \frac{[0 - u]^2}{u\,(1 - u)} \, du + \sum_{k=1}^{n-1} \int_{u_k}^{u_{k+1}} \frac{\left(\frac{k}{n} - u\right)^2}{u\,(1 - u)} \, du + \int_{u_n}^{1} \frac{(1 - u)^2}{u\,(1 - u)} \, du \right\}$$

$$= n \left\{ -u_1 - \ln(1 - u_1) + \sum_{k=1}^{n-1} \int_{u_k}^{u_{k+1}} \frac{\left(\frac{k^2}{n^2} - \frac{2k}{n} u + u^2\right)}{u\,(1 - u)} \, du \right.$$

$$\left. + \left[-1 + u_n - \ln u_n\right] \right\}$$

$$= n\left\{ - u_1 - \ln(1-u_1) + \sum_{k=1}^{n-1} \left[\frac{k^2}{n^2} \int_{u_k}^{u_{k+1}} \frac{du}{u(1-u)} \right] \right.$$

$$+ \sum_{k=1}^{n-1} \left[- \frac{2k}{n} \int_{u_k}^{u_{k+1}} \frac{u}{u(1-u)}\, du \right]$$

$$+ \sum_{k=1}^{n-1} \left[\int_{u_k}^{u_{k+1}} \frac{u^2}{u(1-u)}\, du \right]$$

$$\left. + \left[-1 + u_n - \ln u_n \right] \right\}$$

$$= n\left\{ - u_1 - \ln(1-u_1) \right.$$

$$+ \sum_{k=1}^{n-1} \frac{k^2}{n^2} \left[\ln(u_{k+1}) - \ln(1-u_{k+1}) - \ln u_k + \ln(1-u_k) \right]$$

$$+ \sum_{k=1}^{n-1} \frac{2k}{n} \left[\ln(1-u_{k+1}) - \ln(1-u_k) \right]$$

$$+ \sum_{k=1}^{n-1} \left[- u_{k+1} - \ln(1-u_{k+1}) + u_k + \ln(1-u_k) \right]$$

$$\left. + \left[-1 + u_n - \ln u_n \right] \right\}$$

$$= n\left\{ -u_1 - \ln(1-u_1) + \sum_{k=1}^{n-1} -\left(\frac{2k-1}{n^2}\right)\left[\ln u_k - \ln(1-u_k)\right]\right.$$

$$+ \left(\frac{n-1}{n}\right)^2 \left[\ln u_n - \ln(1-u_n)\right]$$

$$+ \sum_{k=1}^{n-1}\left[-\frac{2}{n}\ln(1-u_k)\right]$$

$$+ \left(\frac{2n-2}{n}\right)\ln(1-u_n)$$

$$+ u_1 + \ln(1-u_1) - u_n - \ln(1-u_n)$$

$$\left. + \left[-1 + u_n - \ln u_n\right]\right\} .$$

$$= n\left\{ -1 - \sum_{k=1}^{n}\left[\left(\frac{2k-1}{n^2}\right)\ln u_k + \left(\frac{2(n-k)+1}{n^2}\right)\ln(1-u_k)\right]\right\}$$

$$= -n - \frac{1}{n}\sum_{k=1}^{n}\left[(2k-1)\ln u_k + (2(n-k)+1)\ln(1-u_k)\right] .$$

INTENTIONALLY LEFT BLANK.

APPENDIX C: DERIVATION OF THE CRAMÉR-VON MISES STATISTIC

INTENTIONALLY LEFT BLANK.

APPENDIX C: DERIVATION OF THE CRAMÉR-VON MISES STATISTIC

For an ordered sample $x_1 \leq x_2 \leq \cdots \leq x_n$ the empirical distribution function is defined as

$$F_n(x) = \begin{cases} 0 & x < x_1 \\ \dfrac{k}{n} \quad \text{for} & x_k \leq x < x_{k+1} \\ 1 & x_n \leq x \end{cases} .$$

The Cramér-von Mises statistic may be written

$$n \int_{-\infty}^{\infty} [F_n(x) - F(x)]^2 dF(x)$$

$$= n \left\{ \int_{-\infty}^{x_1} [F_n(x) - F(x)]^2 \, dF(x) + \sum_{k=1}^{n-1} \int_{x_k}^{x_{k+1}} [F_n(x) - F(x)]^2 \, dF(x) \right.$$

$$\left. + \int_{x_n}^{\infty} [F_n(x) - F(x)]^2 \, dF(x) \right\}$$

Let $F(x) = u$
$dF(x) = du$
$F(x_i) = u_i$

Then

$$n \int_{-\infty}^{\infty} [F_n(x) - F(x)]^2 dF(x)$$

$$= n \left\{ \int_0^{u_1} \left[0 - u\right]^2 du + \sum_{k=1}^{n-1} \int_{u_k}^{u_{k+1}} \left[\frac{k}{n} - u\right]^2 du + \int_{u_n}^{1} \left[1 - u\right]^2 du \right\}$$

$$= n \left\{ \frac{1}{3} u_1{}^3 + \sum_{k=1}^{n-1} \left[\frac{k^2}{n^2} u - \frac{k}{n} u^2 + \frac{1}{3} u^3\right]_{u_k}^{u_{k+1}} + \left[\frac{1}{3} - u_n + u_n^2 - \frac{1}{3} u_n^3\right] \right\}$$

$$= n \left\{ \frac{1}{3} u_1^3 + \frac{1}{n^2} \sum_{k=1}^{n-1} k^2 (u_{k+1} - u_k) - \frac{1}{n} \sum_{k=1}^{n-1} k\,(u_{k+1}^2 - u_k^2) \right.$$

$$\left. + \frac{1}{3} \sum_{k=1}^{n-1} (u_{k+1}^3 - u_k^3) + \left[\frac{1}{3} - u_n + u_n^2 - \frac{1}{3} u_n^3\right] \right\}$$

$$= n \left\{ \frac{1}{3} u_1^3 + \frac{1}{n^2} \left[\sum_{k=1}^{n-1} (-(2k-1) u_k) + (n-1)^2 u_n\right] \right.$$

$$- \frac{1}{n} \left[- \sum_{k=1}^{n-1} u_k^2 + (n-1) u_n^2\right]$$

$$\left. + \frac{1}{3} u_n^3 - \frac{1}{3} u_1^3 + \left[\frac{1}{3} - u_n + u_n^2 - \frac{1}{3} u_n^3\right] \right\}$$

$$= n \left\{ \sum_{k=1}^{n-1} \left[\frac{1}{n} u_k^2 - \left(\frac{2k-1}{n^2}\right) u_k\right] + \frac{1}{3} + \frac{1}{n} u_n^2 - \left(\frac{2n-1}{n^2}\right) u_n \right\}$$

$$= \sum_{k=1}^{n} \left[u_k^2 - \left(\frac{2k-1}{n}\right) u_k\right] + \frac{n}{3}$$

Completing the square,

$$= \sum_{k=1}^{n} \left[u_k - \left(\frac{2k-1}{2n} \right) \right]^2 + \frac{n}{3} - \sum_{k=1}^{n} \left(\frac{2k-1}{2n} \right)^2$$

$$= \sum_{k=1}^{n} \left[u_k - \left(\frac{2k-1}{2n} \right) \right]^2 + \frac{n}{3} - \frac{1}{4n^2} \left(1^2 + 3^2 + \cdots + (2n-1)^2 \right)$$

$$= \sum_{k=1}^{n} \left[u_k - \left(\frac{2k-1}{2n} \right) \right]^2 + \frac{n}{3} - \frac{1}{4n^2} \left(\frac{(n)(4n^2-1)}{3} \right)$$

$$= \sum_{k=1}^{n} \left[u_k - \left(\frac{2k-1}{2n} \right) \right]^2 + \frac{1}{12n} .$$

INTENTIONALLY LEFT BLANK.

No of Copies	Organization
(Unclass., unlimited) 12 (Unclass., limited) 2 (Classified) 2	Administrator Defense Technical Info Center ATTN: DTIC-DDA Cameron Station Alexandria, VA 22304-6145
1	HQDA (SARD-TR) WASH DC 20310-0001
1	Commander US Army Materiel Command ATTN: AMCDRA-ST 5001 Eisenhower Avenue Alexandria, VA 22333-0001
1	Commander US Army Laboratory Command ATTN: AMSLC-DL Adelphi, MD 20783-1145
2	Commander Armament RD&E Center US Army AMCCOM ATTN: SMCAR-MSI Picatinny Arsenal, NJ 07806-5000
2	Commander Armament RD&E Center US Army AMCCOM ATTN: SMCAR-TDC Picatinny Arsenal, NJ 07806-5000
1	Director Benet Weapons Laboratory Armament RD&E Center US Army AMCCOM ATTN: SMCAR-LCB-TL Watervliet, NY 12189-4050
1	Commander US Army Armament, Munitions and Chemical Command ATTN: SMCAR-ESP-L Rock Island, IL 61299-5000
1	Commander US Army Aviation Systems Command ATTN: AMSAV-DACL 4300 Goodfellow Blvd. St. Louis, MO 63120-1798
1	Director US Army Aviation Research and Technology Activity Ames Research Center Moffett Field, CA 94035-1099

No of Copies	Organization
1	Commander US Army Missile Command ATTN: AMSMI-RD-CS-R (DOC) Redstone Arsenal, AL 35898-5010
1	Commander US Army Tank Automotive Command ATTN: AMSTA-TSL (Technical Library) Warren, MI 48397-5000
1	Director US Army TRADOC Analysis Command ATTN: ATAA-SL White Sands Missile Range, NM 88002-5502
(Class. only) 1	Commandant US Army Infantry School ATTN: ATSH-CD (Security Mgr.) Fort Benning, GA 31905-5660
(Unclass. only) 1	Commandant US Army Infantry School ATTN: ATSH-CD-CSO-OR Fort Benning, GA 31905-5660
(Class. only) 1	The Rand Corporation P.O Box 2138 Santa Monica, CA 90401-2138
1	Air Force Armament Laboratory ATTN: AFATL/DLODL Eglin AFB, FL 32542-5000
	<u>Aberdeen Proving Ground</u> Dir, USAMSAA ATTN: AMXSY-D AMXSY-MP, H. Cohen Cdr, USATECOM ATTN: AMSTE-TO-F Cdr, CRDEC, AMCCOM ATTN: SMCCR-RSP-A SMCCR-MU SMCCR-MSI Dir, VLAMO ATTN: AMSLC-VL-D

No of Copies	Organization
1	US Army Concepts Analysis Agency ATTN: Carl B. Bates 8120 Woodmont Avenue Bethesda, MD 20814-2797
	Aberdeen Proving Ground Dir, USAMSAA ATTN: AMXSY-RI, H. Pasini AMXSY-CT, A. Benton AMXSY-R, B. Bennett AMXSY-R, D. O'Neill AMXSY-R, W. Mowchan AMSXY-R, R. Mioduski Cdr, USATECOM ATTN: AMSTE-TA-G, N. Dunn AMSTE-TA-G, L. West Cdr, CRDEC, AMCCOM ATTN: SMCCR-RSP-C, L. Sturdivan Cdr, USACSTA ATTN: STECS-DA-PS, V. Visnaw

USER EVALUATION SHEET/CHANGE OF ADDRESS

This laboratory undertakes a continuing effort to improve the quality of the reports it publishes. Your comments/answers below will aid us in our efforts.

1. Does this report satisfy a need? (Comment on purpose, related project, or other area of interest for which the report will be used.) ______________________________

2. How, specifically, is the report being used? (Information source, design data, procedure, source of ideas, etc.) ______________________________

3. Has the information in this report led to any quantitative savings as far as man-hours or dollars saved, operating costs avoided, or efficiencies achieved, etc? If so, please elaborate. ______________________________

4. General Comments. What do you think should be changed to improve future reports? (Indicate changes to organization, technical content, format, etc.) ______________________________

BRL Report Number ____________ Division Symbol ____________

Check here if desire to be removed from distribution list. ____

Check here for address change. ____

Current address: Organization ______________________________
Address ______________________________

---------------------------FOLD AND TAPE CLOSED---------------------------

Director
U.S. Army Ballistic Research Laboratory
ATTN: SLCBR-DD-T
Aberdeen Proving Ground, MD 21005-5066

OFFICIAL BUSINESS
PENALTY FOR PRIVATE USE $300

NO POSTAGE NECESSARY IF MAILED IN THE UNITED STATES

BUSINESS REPLY LABEL
FIRST CLASS PERMIT NO. 12062 WASHINGTON D.C.

POSTAGE WILL BE PAID BY DEPARTMENT OF THE ARMY

Director
U.S. Army Ballistic Research Laboratory
ATTN: SLCBR-DD-T
Aberdeen Proving Ground, MD 21005-9989